

STIC-Biotech/ChemLib

95090

From: Rao, Manjunath N.
Sent: Tuesday, May 27, 2003 3:48 PM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 09/980771

From: Manjunath N. Rao
Art Unit 1652, Room 10A11
Mail Box in Room 10D 01
Phone: 306-5681

Date: 5-27-03

Please search the following as soon as possible for application with serial number **09/980771**

SEQ ID NO: 2, 4, 6, 8 against all commercial nucleic acid databases including issued patents database and pending application database and provide a print of all results.

SEQ ID NO:3, and amino acids 58-708 of SEQ ID NO:3, SEQ ID NO:5, 7, 9 against all commercial protein databases including issued patents database and pending application database and provide a print of all results.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.
Biotechnology Patent Examiner
Art Unit 1652, Room 10A11
Mail Box in 10D01
Crystal Mall 1, USPTO.

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 5/28
Date Completed: 6/10
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

This page Blank (uspto)

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 08:28:25 ; Search time 3691.76 Seconds
(without alignments)
16743.883 Million cell updates/sec

Title: US-09-980-771A-2

Perfect score: 2124

Sequence: 1 atgcgtgtgcctctaccag.....ccgccaccgcgaagtcgcgc 2124

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pin.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2124	100.0	2124	6	AX049323	Sequence
2	2124	100.0	3117	6	AX049322	Sequence
3	2122.4	99.9	3240	8	AF026420	Chlamydom
4	1953	91.9	1953	6	AX049325	Sequence
5	1593	75.0	1593	6	AX049329	Sequence
6	1314	61.9	1314	6	AX049327	Sequence
7	962.2	45.3	5856	8	AF433156	Chlamydom
8	537	25.3	2289	8	AF286320	Triticum
9	531.2	25.0	1812	8	AF486514	Hordeum v
10	531.2	25.0	2311	8	HWAXYR	AF07932 Barley mRNA
11	529.6	24.9	1827	8	AF486515	Hordeum v
12	529.6	24.9	1827	8	AF486518	Hordeum v
13	527.4	24.8	2127	8	AF163319	Triticum
14	526.4	24.8	1812	8	AF486519	Hordeum v
15	526.4	24.8	1827	8	AF486516	Hordeum v
16	526.4	24.8	1827	8	AF486517	Hordeum v
17	525.8	24.8	1605	8	AF250137	Triticum
18	507.2	23.9	2028	8	AF113844	Triticum
19	507	23.9	2186	8	TAWAXYSS	X57233 Wheat waxy
20	507	23.9	2267	8	AY050174	Triticum
21	481.4	22.7	1801	8	AF113843	Triticum
22	478	22.5	2267	6	AR195560	Sequence
23	475.4	22.4	2115	8	AB089141	Setaria i
24	457.4	21.5	2287	8	SBU23945	U23945 Sorghum bic
25	449.2	21.1	1937	8	AB066093	Oryza sat
26	449.2	21.1	1937	8	AB066094	Oryza sat
27	447.6	21.1	2542	6	AR106491	Sequence
28	447.6	21.1	2542	8	OSWX	X62134 O. sativa wa
29	347	16.3	2168	8	MEGBSS	X74160 M.esculenta
30	344.8	16.2	2270	8	AF210699	Perilla f
31	332.6	15.7	2221	8	AB029546	Phaseolus
32	321.6	15.1	2211	8	AB071604	Ipomoea b
33	316.8	14.9	2161	6	AX394246	Sequence
34	313.8	14.8	3049	8	AF026421	Chlamydom
35	313.2	14.7	2345	8	AMA6293	Antirrhin
36	305.6	14.4	2124	8	IBU44126	U44126 Ipomoea bat
37	298	14.0	2092	8	AY094405	Arabidops
38	298	14.0	2115	8	AY088544	Arabidops
39	298	14.0	2192	8	AY123983	Arabidops
40	289.6	13.6	2107	8	AF097922	Astragalu
41	283.2	13.3	764	8	AF446083	Miscanthu
42	280.2	13.2	2035	8	PSSTARSYN	X88789 P. sativum m
43	274.2	12.9	2081	8	AF109395	Triticum
44	274	12.9	764	8	AF446082	Microsteg
45	273.2	12.9	721	8	AF446081	Cleistach

ALIGNMENTS

RESULT 1
AX049323
LOCUS AX049323
DEFINITION Sequence 2 from Patent WO0071734.
ACCESSION AX049323
VERSION AX049323.1 GI:12226090
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 2124)
AUTHORS D'Hulst,C. and Ball,S
TITLE Starch granules containing a recombinant polypeptide of interest,
method for obtaining same and uses
JOURNAL Patent: WO 0071734-A 2 30-NOV-2000;

Query Match	100.0%;	Score 2124;	DB 6;	Length 2124;
Best Local Similarity	100.0%;	Pred. No. 7.le-202;		
Matches 2124;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	ATGGCTGTGGCTCTACAGCGCCCGCAGCAGCGCGGTCTCTATCGTCATCAATGCCGCG	60		
Db 1	ATGGCTGTGGCTCTACAGCGCCCGCAGCAGCGCGGTCTCTATCGTCATCAATGCCGCG	60		
Qy 61	TCGTCGGGTGCAAGAGACCGCAACACAGTCTGCTGGTGAGCTTGCTCGTGGCTCGCA	120		
Db 61	TCGTCGGGTGCAAGAGACCGCAACACAGTCTGCTGGTGAGCTTGCTCGTGGCTCGCA	120		
Qy 121	CGAAGTCCACTCGCGTTCGGCTGTACTGGTGGCCACTGGTGCCACTTGGCGCTGGAC	180		
Db 121	CGAAGTCCACTCGCGTTCGGCTGTACTGGTGGCCACTGGTGCCACTTGGCGCTGGAC	180		
Qy 181	ATCGTGATGGTGTGCTGTAGGTGCGCCCTGGTGTCCAAAGACGGCGGCGCTGGCGGATGTG	240		
Db 181	ATCGTGATGGTGTGCTGTAGGTGCGCCCTGGTGTCCAAAGACGGCGGCGCTGGCGGATGTG	240		
Qy 241	ACTGGTGGCCTGCCTATTGAGCTGGTCAAGCGCGGCCACCGCGTCATGACATTCGCCCT	300		
Db 241	ACTGGTGGCCTGCCTATTGAGCTGGTCAAGCGCGGCCACCGCGTCATGACATTCGCCCT	300		
Qy 301	CGCTACGACCACTACGCTGAGCCTGGGACACCTCGGTGTCTGGACATCATGGGGGAG	360		
Db 301	CGCTACGACCACTACGCTGAGCCTGGGACACCTCGGTGTCTGGACATCATGGGGGAG	360		
Qy 361	AAGTTCGCTACTTCCACTCCATCAAGAAGGGGTGCACCGCGTGTGGATTGACCAACCC	420		
Db 361	AAGTTCGCTACTTCCACTCCATCAAGAAGGGGTGCACCGCGTGTGGATTGACCAACCC	420		
Qy 421	TGGTTCCTGGCCAAAGTCTGGGGCAAGACCGGCTCCAAAGTGTACGGGCCCGCTCGGC	480		
Db 421	TGGTTCCTGGCCAAAGTCTGGGGCAAGACCGGCTCCAAAGTGTACGGGCCCGCTCGGC	480		
Qy 481	GCTGACTACTTGGACAAACCAAGGGCTTCGCCCTGTCTCAAGGGCGCTATTGAGGCT	540		
Db 481	GCTGACTACTTGGACAAACCAAGGGCTTCGCCCTGTCTCAAGGGCGCTATTGAGGCT	540		
Qy 541	GCOCGCGTGTGCCCTTCGGCCCGCGGAGGACTGCGCTTCTCGTGGCCAAAGTGGCAC	600		
Db 541	GCOCGCGTGTGCCCTTCGGCCCGCGGAGGACTGCGCTTCTCGTGGCCAAAGTGGCAC	600		


```
Db 1681 CTGGAGGAGTGGTGTACGGCAAGGCGCGCTGGCCACCGCCACAGAGAGAGATCAAG 1740
QY 1741 GTGCCCGTTCGCGAGAGATCCCGGCGACCTGCGCGCGGTGTCTTACGCCGCCCAACACC 1800
Db 1741 GTGCCCGTTCGCGAGAGATCCCGGCGACCTGCGCGCGGTGTCTTACGCCGCCCAACACC 1800
QY 1801 CTGAAGCCCGTGTCCGCGCTCGGTGAGGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCG 1860
Db 1801 CTGAAGCCCGTGTCCGCGCTCGGTGAGGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCG 1860
QY 1861 ACCGCCCGCGCATGGGCGCGTGGCGCGGACCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 1920
Db 1861 ACCGCCCGCGCATGGGCGCGTGGCGCGGACCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 1920
QY 1921 GCCACGCCAAGGTGACCACTTACAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1980
Db 1921 GCCACGCCAAGGTGACCACTTACAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1980
QY 1981 GCTGCGCTCAAGCTGCGCGGTGAGGCGTCCACCACTTCCACCACTTCCAGCAAGCGCGTGC 2040
Db 1981 GCTGCGCTCAAGCTGCGCGGTGAGGCGTCCACCACTTCCACCACTTCCAGCAAGCGCGTGC 2040
QY 2041 TCCACGGCAACGGCAAGCGTGCCTCGCGCTCCAAAGACTTGGGCTGCCAAGCGCGCGTGTTC 2100
Db 2041 TCCACGGCAACGGCAAGCGTGCCTCGCGCTCCAAAGACTTGGGCTGCCAAGCGCGCGTGTTC 2100
QY 2101 TCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2124
Db 2101 TCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2124

RESULT 2
AX049322
LOCUS AX049322 3117 bp DNA linear PAT 12-JAN-2001
DEFINITION Sequence 1 from Patent WO0071734.
ACCESSION AX049322
VERSION AX049322.1 GI:12226089
KEYWORDS
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE 1 (bases 1 to 3117)
AUTHORS D'Hulst,C. and Ball,S.
TITLE Starch granules containing a recombinant polypeptide of interest,
METHOD for obtaining same and uses
PATENT: WO 0071734-A 1 30-NOV-2000;
JOURNAL CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES
source
1..3117
/organism="Chlamydomonas reinhardtii"
BASE COUNT 599 a 994 c 1011 g 513 t
ORIGIN

Query Match 100.0%; Score 2124; DB 6; Length 3117;
Best Local Similarity 100.0%; Pred. No. 6.3e-202;
Matches 2124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGTTCCTTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 15 ATGCTGTTCCTTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 74
QY 61 TCGTTCGGTGTCAAGAGACCGGCAACAGCTGCTGCTGAGCTTGTCTGCTGCGTCCGCA 120
Db 75 TCGTTCGGTGTCAAGAGACCGGCAACAGCTGCTGCTGAGCTTGTCTGCTGCGTCCGCA 134
QY 121 CGAAGTCCACCTCGCGCTCGGTGTTACTGGTCCACTGTGCGCACTTGGCGCTGGAC 180
Db 135 CGAAGTCCACCTCGCGCTCGGTGTTACTGGTCCACTGTGCGCACTTGGCGCTGGAC 194
QY 181 ATCGTGAATGTTGCTGTGAGTGCGCCCTTGGTCCAAAGACGGCGCGCGCGCGCGCGCG 240
```

```
Db 195 ATCGTGAATGTTGCTGTGAGTTCGCCCTTGGTCCAAAGACGGCGCGCGCGCGCGCGATG 254
QY 241 ACTGGTGGCTGCGATTGAGCTGCTCAAGCGCGGCCACCGCGTCAATGACCAATTGCCCT 300
Db 255 ACTGGTGGCTGCGATTGAGCTGCTCAAGCGCGGCCACCGCGTCAATGACCAATTGCCCT 314
QY 301 CGTACGACCACTGACGCTGACGCTGGGACACCTCGGTGCTGCTGACATCATGCGCGAG 360
Db 315 CGTACGACCACTGACGCTGACGCTGGGACACCTCGGTGCTGCTGACATCATGCGCGAG 374
QY 361 AAGTTCGGTACTTCCACTTCCATCAAGAAGGGCGTCAACCGCTTACGGCCCCCGCTCCGC 420
Db 375 AAGTTCGGTACTTCCACTTCCATCAAGAAGGGCGTCAACCGCTTACGGCCCCCGCTCCGC 434
QY 421 TGGTTCCTGGCCAAAGTCTGGGCAAGACCGGCTCCAAGCTGACGGCCCCCGCTCCGC 480
Db 435 TGGTTCCTGGCCAAAGTCTGGGCAAGACCGGCTCCAAGCTGACGGCCCCCGCTCCGC 494
QY 481 GCTGACTACTTGGCAACCAACGAGCTTCGCCCTTCTGCAAGGCCGCTATTGAGGCT 540
Db 495 GCTGACTACTTGGCAACCAACGAGCTTCGCCCTTCTGCAAGGCCGCTATTGAGGCT 554
QY 541 GCCCGGTGCTGCCCTTGGCCCCCGGAGGACTGCGTCTTCTGCGCCAAACGACTGGCAC 600
Db 555 GCCCGGTGCTGCCCTTGGCCCCCGGAGGACTGCGTCTTCTGCGCCAAACGACTGGCAC 614
QY 601 TCCGCCCTGGTCCCGCTCTGCTGAAGACGAGTACCGAGCCCAAGGGCGAGTTCACCAAG 660
Db 615 TCCGCCCTGGTCCCGCTCTGCTGAAGACGAGTACCGAGCCCAAGGGCGAGTTCACCAAG 674
QY 661 GCCAAGTCTGCTGCTTATCCAAACATCGCTTCCAGGGCGCATGTGGGAGGAGCT 720
Db 675 GCCAAGTCTGCTGCTTATCCAAACATCGCTTCCAGGGCGCATGTGGGAGGAGCT 734
QY 721 TTCAAGACACGAGCTGCCCGAGCGCGCTTGAAGCTGCGCTTCTCGACAGGCTAT 780
Db 735 TTCAAGACACGAGCTGCCCGAGCGCGCTTGAAGCTGCGCTTCTCGACAGGCTAT 794
QY 781 GCCAAGTCTTACCTGAGGCGCACCCCTTGAAGGAGGAGGAGAGAGCCCGCTGACGGGA 840
Db 795 GCCAAGTCTTACCTGAGGCGCACCCCTTGAAGGAGGAGGAGAGAGCCCGCTGACGGGA 854
QY 841 AAGACTTACAAGAGATCAACTGGCTGAAGGTGCATTATCGCGCGCGCAAGGTGTG 900
Db 855 AAGACTTACAAGAGATCAACTGGCTGAAGGTGCATTATCGCGCGCGCAAGGTGTG 914
QY 901 ACTGTGCGCGCACTACCGACGAGATCGCTCCGATGCGCGCGCGGTGTGAGGTG 960
Db 915 ACTGTGCGCGCACTACCGACGAGATCGCTCCGATGCGCGCGCGGTGTGAGGTG 974
QY 961 GACACCGCTATCCGCGCAAGGCGATTGAGGCGATTGAGGCGATGAGGCGATGAGGAG 1020
Db 975 GACACCGCTATCCGCGCAAGGCGATTGAGGCGATTGAGGCGATGAGGCGATGAGGAG 1034
QY 1021 TGAACCCCAAGACCGCAAGTTCCTGCTCGCCCTACGACCAAGAGAGAGAGTACGCC 1080
Db 1035 TGAACCCCAAGACCGCAAGTTCCTGCTCGCCCTACGACCAAGAGAGAGAGTACGCC 1094
QY 1081 GGAAGGCCCGCGCAAGGAGGCGCTTGCAGCCGAGCTGGCGCTGCTGTGGAGACCCAC 1140
Db 1095 GGAAGGCCCGCGCAAGGAGGCGCTTGCAGCCGAGCTGGCGCTGCTGTGGAGACCCAC 1154
QY 1141 GCCCGCTGTGCGCTTATCGCGCGCGCTGAGGAGGAGGAGGAGTGGACATCATCTG 1200
Db 1155 GCCCGCTGTGCGCTTATCGCGCGCGCTGAGGAGGAGGAGGAGTGGACATCATCTG 1214
QY 1201 GCGCGCTGCGCAAGATCTTGGCCACCGCCCAAGGTGAGATGCGCATCTCTGGGTACCGC 1260
Db 1215 GCGCGCTGCGCAAGATCTTGGCCACCGCCCAAGGTGAGATGCGCATCTCTGGGTACCGC 1274
QY 1261 AAGGCCCTTACGAGAGCTGTGTAACCGCATCGGCACCAAGTACAAGGCGCGCGCAAG 1320
```

```
Db 1275 AAGCGCGCTACGAGAAGCTGGTGAACGCCATCGGCACCAAGTACAAAGGCGCGCAAG 1334
QY 1321 GGCGTGTCAAGTCTTCGCGCGCCCTTGGCGGCACATGCTACCGCGCGCGGCTTCAATG 1380
Db 1335 GGCGTGTCAAGTCTTCGCGCGCCCTTGGCGGCACATGCTACCGCGCGCGGCTTCAATG 1394
QY 1381 CTGCTGCCCTTCGCGCTTCGAGAGCCTTGGCGGCTGATCCAGCTGACGCGCATGCACTACGGT 1440
Db 1395 CTGCTGCCCTTCGCGCTTCGAGAGCCTTGGCGGCTGATCCAGCTGACGCGCATGCACTACGGT 1454
QY 1441 ACCGTGCCCTTCGCTTCGAGAGCCTTGGCGGCTGATCCAGCTGACGCGCATGCACTACGGT 1500
Db 1455 ACCGTGCCCTTCGCTTCGAGAGCCTTGGCGGCTGATCCAGCTGACGCGCATGCACTACGGT 1514
QY 1501 GGCTTCCACATGGCGCGCTTGAACCCCGACAGCTGACGAGCTGACGCGCGCGCTG 1560
Db 1515 GGCTTCCACATGGCGCGCTTGAACCCCGACAGCTGACGAGCTGACGCGCGCGCTG 1574
QY 1561 GCCGCCACCGTGGCGGCTGACGAGCTGTTTGGCGGCGCGGCTTACCGCGAGATGGT 1620
Db 1575 GCCGCCACCGTGGCGGCTGACGAGCTGTTTGGCGGCGCGGCTTACCGCGAGATGGT 1634
QY 1621 GCCAACTGATCAGCAGGAGCTGCTGCTGTCGAAAGCCCGCCCAAGAGTGGGAGGCGCTG 1680
Db 1635 GCCAACTGATCAGCAGGAGCTGCTGCTGTCGAAAGCCCGCCCAAGAGTGGGAGGCGCTG 1694
QY 1681 CTGAGGAGGTGGTGTACGGCAAGGGCGGCTGGCCACCGCAAGAGGAGGATCAAG 1740
Db 1695 CTGAGGAGGTGGTGTACGGCAAGGGCGGCTGGCCACCGCAAGAGGAGGATCAAG 1754
QY 1741 GTGCCCGTTCGCGAAGATATCCCGCGGACCTGCTGCGCGCGGCTTACGCGCCCAACACC 1800
Db 1755 GTGCCCGTTCGCGAAGATATCCCGCGGACCTGCTGCGCGCGGCTTACGCGCCCAACACC 1814
QY 1801 CTGAAGCCCGTTCGCGTTCGAGGGCAAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1860
Db 1815 CTGAAGCCCGTTCGCGTTCGAGGGCAAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1874
QY 1861 ACCGCGCGCGCATGGCGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1920
Db 1875 ACCGCGCGCGCATGGCGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1934
QY 1921 GCCACCCCGCAAGGTGACCATACGAGCCCGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1980
Db 1935 GCCACCCCGCAAGGTGACCATACGAGCCCGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1994
QY 1981 GCTGCGCTCAAGCTGGCGGCTGAGCGCTCCACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2040
Db 1995 GCTGCGCTCAAGCTGGCGGCTGAGCGCTCCACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2054
QY 2041 TCCACGCGCAAGCGGCTGCGCTGCGGCTTCCAGAGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGG 2100
Db 2055 TCCACGCGCAAGCGGCTGCGCTGCGGCTTCCAGAGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGG 2114
QY 2101 TCCGCGCGCGCGCAAGTTCGCGG 2124
Db 2115 TCCGCGCGCGCGCAAGTTCGCGG 2138

RESULT 3
AF026420 3240 bp mRNA linear PLN 05-NOV-2001
LOCUS Chlamydomonas reinhardtii granule-bound starch synthase I precursor
DEFINITION (STA2) mRNA, complete cds.
ACCESSION AF026420
VERSION AF026420.3 GI:16716334
KEYWORDS Chlamydomonas reinhardtii.
SOURCE Chlamydomonas reinhardtii
ORGANISM Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
REFERENCE 1 (bases 1 to 3240)
AUTHORS D'Hulst,C., Wattedbled,F., Ral,J.-P., Abel,G.J., Kossmann,J. and
```



```
QY 1672 GAGGGCTGCTGGAGGAGTGGTGTACGGCAAGGGCGGGCGGCGGCGCCACCGCCCAAGAAGGAG 1731
|||||
Db 1501 GAGGGCTGCTGGAGGAGTGGTGTACGGCAAGGGCGGGCGGCGGCGCCACCGCCCAAGAAGGAG 1560
|||||
QY 1732 GAGATCAAGGTGCCCGTTGCGGAGAAATCCCGGGGACCTGCGCGCGGTGTCTACGCC 1791
|||||
Db 1561 GAGATCAAGGTGCCCGTTGCGGAGAAATCCCGGGGACCTGCGCGCGGTGTCTACGCC 1620
|||||
QY 1792 CCCAACACCGTGAAGCCCGTGTCCGCTCCGTGGAGGCAACGGCGCGCGCCCAAG 1851
|||||
Db 1621 CCCAACACCGTGAAGCCCGTGTCCGCTCCGTGGAGGCAACGGCGCGCGCCCAAG 1680
|||||
QY 1852 GTCGGCACCGCCCGCCATGGGCGCGTGGCGCGGACCCCGCTCGGGCCCTCG 1911
|||||
Db 1681 GTCGGCACCGCCCGCCATGGGCGCGTGGCGCGGACCCCGCTCGGGCCCTCG 1740
|||||
QY 1912 CCCGCCCGCGCCACCCCAAGGTGACCACTTACAAGCCGCGCCCTGCCCGCCACCGCCCAAG 1971
|||||
Db 1741 CCCGCCCGCGCCACCCCAAGGTGACCACTTACAAGCCGCGCCCTGCCCGCCACCGCCCAAG 1800
|||||
QY 1972 CCCAAGACCGTGGCTTCAAGCTGGCGGTGAGGCTTCCACCACTCGACCTCGGAGAAC 2031
|||||
Db 1801 CCCAAGACCGTGGCTTCAAGCTGGCGGTGAGGCTTCCACCACTCGGAGAAC 1860
|||||
QY 2032 GGGCTGCCCTCCCAACGCAACGCAACGCTGCTCGGCTCCCAAGACCTCGGCTGCCAAG 2091
|||||
Db 1861 GGGCTGCCCTCCCAACGCAACGCAACGCTGCTCGGCTCCCAAGACCTCGGCTGCCAAG 1920
|||||
QY 2092 CCCTGTGCTCCGCGCCACCCGCAAGTCCGCC 2124
|||||
Db 1921 CCCTGTGCTCCGCGCCACCCGCAAGTCCGCC 1953
|||||

RESULT 5
AX049329
LOCUS AX049329
DEFINITION Sequence 8 from Patent WO0071734.
ACCESSION AX049329
VERSION AX049329.1 GI:12226096
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 1593)
AUTHORS D'Hulst,C. and Ball,S.
TITLE Starch granules containing a recombinant polypeptide of interest,
method for obtaining same and uses
Patent: WO 0071734-A 8 30-NOV-2000;
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES
Location/Qualifiers
1..1593
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="fragment de l'DNAC complet codant la GBSSI de
Chlamydomonas reinhardtii"
<1..>1593
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/db_protein="GI:12226097"
/db_xref="GI:12226097"
/translation="ALDLYVAAEVPWSTGGLGVDVGTGLPIELVKRHRVMTIAPR
YQYADAWDTSVVDLMGKVRFFSHKKGVRHVWIDHPWFLAKVWGTGSKLYGPRS
GADLYDNRRFLAFKAALEAARVLPFGPEPCVFVANDWHSALVPLKDEYQPRGQ
PTKASVLAHNTAFGRWMEAFKDTLPQAFDKLAFSDGVAKVYVTEATPMEDEK
PPLTKYTKIINWKGIIAADKLTVPNSYATEIAADAAGVELDVTIRAKGIEGIV
NGMDIEWNPDKTFLSAPYDONSVYGAARAAKEALQELGVPDPTAPLAFIGLE
EQGVDIILALPK ILATEPKVQIALLGTCKAAEKLIVNALTGYKRGAGVYKFSAPL
AHLRTAGDFMLVPSFEPCGLIQLHAMHYIVPVVASTGGLDVTVKEGVTGFHMGAL
NPDKLEADADALAAVFRASEVFAAGGRYPMPVANCISQDLSWSKPAQRWGLLEEV
YKGGVYATAKKEIKVPVAKIP"
BASE COUNT 290 a 539 c 510 g 254 t
ORIGIN
```

```
Query Match 75.0%; Score 1593; DB 6; Length 1593;
Best Local Similarity 100.0%; Pred. No. 2.8e-149;
Matches 1593; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 GCGTGGACATCGTATGATGGTTGCTGCTGAGGTGCGCCCTTGGTCCAAAGACGGCGGCGCTG 231
Db 1 GCGTGGACATCGTATGATGGTTGCTGCTGAGGTGCGCCCTTGGTCCAAAGACGGCGGCGCTG 60
QY 232 GGCATGTGACTGTGGGCTGCTATTCAGCTGCTCAAGCGCGGCGACCGGTATGACC 291
Db 61 GGCATGTGACTGTGGGCTGCTATTCAGCTGCTCAAGCGCGGCGACCGGTATGACC 120
QY 292 ATTGCCCTCCCTACGACCACTAGCTGACGCTGGGACACTCGGTGGTGGTGGACATC 351
Db 121 ATTGCCCTCCCTACGACCACTAGCTGACGCTGGGACACTCGGTGGTGGTGGACATC 180
QY 352 ATGGGCGAGAGGTCCGCTACTTCCACTCCATCAAGAAGGCGGTGCACCGGTGTGAT 411
Db 181 ATGGGCGAGAGGTCCGCTACTTCCACTCCATCAAGAAGGCGGTGCACCGGTGTGAT 240
QY 412 GACCACCCCTGTTCTGCGCAAGTCTGGGCGAGACCGGCTCCAAGCTGTACGGCCCC 471
Db 241 GACCACCCCTGTTCTGCGCAAGTCTGGGCGAGACCGGCTCCAAGCTGTACGGCCCC 300
QY 472 CGTCCGGCGCTGACTACTGGACAAACCAAGCGCTTCGCCCTTGTGCAAGCGCGCT 531
Db 301 CGTCCGGCGCTGACTACTGGACAAACCAAGCGCTTCGCCCTTGTGCAAGCGCGCT 360
QY 532 ATTGAGGCTGCCCGGTGCTGCTGCTTGGGCCCGCGGAGGACTGCGTCTTGGTGGCCAA 591
Db 361 ATTGAGGCTGCCCGGTGCTGCTGCTTGGGCCCGCGGAGGACTGCGTCTTGGTGGCCAA 420
QY 592 GACTGGCACTCCGCCCTGGTCCGCTCTGCTGAAGGACGAGTACACGCCCAAGGGCCAG 651
Db 421 GACTGGCACTCCGCCCTGGTCCGCTCTGCTGAAGGACGAGTACACGCCCAAGGGCCAG 480
QY 652 TTCACCAAGGCGCAAGTGGTGTGCTGCTATCCAAACATCGCTTCCAGGGCGCATGTGG 711
Db 481 TTCACCAAGGCGCAAGTGGTGTGCTGCTATCCAAACATCGCTTCCAGGGCGCATGTGG 540
QY 712 GAGGAGCTTTCAGGACACCAAGCTGCCCGCAGCGCGCTTTGACAGCTGGCTTCTCG 771
Db 541 GAGGAGCTTTCAGGACACCAAGCTGCCCGCAGCGCGCTTTGACAGCTGGCTTCTCG 600
QY 772 GACGGCTATGCAAGGTTTACACTGAGGCCACCCCGCATGGAGGAGGACGAGAAGCCCGC 831
Db 601 GACGGCTATGCAAGGTTTACACTGAGGCCACCCCGCATGGAGGAGGACGAGAAGCCCGC 660
QY 832 CTGACGGGAAGACCTACAAGAATCAACTGGGTGAAGGTGGCATTTATCGCCGCGAC 891
Db 661 CTGACGGGAAGACCTACAAGAATCAACTGGGTGAAGGTGGCATTTATCGCCGCGAC 720
QY 892 AAGCTGTGACTGTGCTGCCCACTACGCGACCGAGATCGCTGCGGATCGCGCGGGGT 951
Db 721 AAGCTGTGACTGTGCTGCCCACTACGCGACCGAGATCGCTGCGGATCGCGCGGGGT 780
QY 952 GTGAGCTGGACACCGCTCATCCGCGCAAGGGCATTTGAGGGCATTTGAAAGGCGATGGAC 1011
Db 781 GTGAGCTGGACACCGCTCATCCGCGCGCAAGGGCATTTGAGGGCATTTGAAAGGCGATGGAC 840
QY 1012 ATTGAGGAGTGGAAACCCCAAGACGCAAGTTCCTGTCTCGCCCTACGACCAAGACAGC 1071
Db 841 ATTGAGGAGTGGAAACCCCAAGACGCAAGTTCCTGTCTCGCCCTACGACCAAGACAGC 900
QY 1072 GTCTACCGCGCAGGCGCGCGCAAGGAGGCGCTGACGCGCGGAGCTGGGCTGCGCTGTG 1131
Db 901 GTCTACCGCGCAGGCGCGCGCAAGGAGGCGCTGACGCGCGGAGCTGGGCTGCGCTGTG 960
QY 1132 GACCCACCGCCCGCTTTCGCTTTCATCGCGCGCTGGAGGAGCAAGGGTGTGAC 1191
Db 961 GACCCACCGCCCGCTTTCGCTTTCATCGCGCGCTGGAGGAGCAAGGGTGTGAC 1020
```



```

Db 901 GTCTACGCGGCAAGCGCGCGCCCAAGAGGAGGCGCTTCAGGCGCGAGCTGGCGCTGCTGTG 960
QY 1132 GACCCACGCGCCCTCTGTCCTTCATCGCGCGCCCTGGAGGACGAGAAGGCTGTGAC 1191
Db 961 GACCCACGCGCGCCCTCTGTCCTTCATCGCGCGCCCTGGAGGACGAGAAGGCTGTGAC 1020
QY 1192 ATCACTCTGGCGCGCGCTGCGCAAGATCTGCGCCACCCCAAGGTGCAGATCGCATCTG 1251
Db 1021 ATCACTCTGGCGCGCGCTGCGCAAGATCTGCGCCACCCCAAGGTGCAGATCGCATCTG 1080
QY 1252 GGTACCGCAAGGCGCGCTACGAGAAGCTGTGACGCGCATCGGACCAAGTACAAAGGC 1311
Db 1081 GGTACCGCAAGGCGCGCTACGAGAAGCTGTGACGCGCATCGGACCAAGTACAAAGGC 1140
QY 1312 CGCGCAAGGCGCGCTGCAAGTCTGCGCGCCCTGGCGGCACATGCTCACCGCGCGCC 1371
Db 1141 CGCGCAAGGCGCGCTGCAAGTCTGCGCGCCCTGGCGGCACATGCTCACCGCGCGCC 1200
QY 1372 GACTTCATGCTGCGCGCGCTGCGAGCCCTGCGGCGCTGATCCAGCTGCAGCCCATG 1431
Db 1201 GACTTCATGCTGCGCGCGCTGCGAGCCCTGCGGCGCTGATCCAGCTGCAGCCCATG 1260
QY 1432 CACTACGCTACGCGCGCTGCTGAGCTCCACCGCGCGCTGTCGACACCGTC 1485
Db 1261 CACTACGCTACGCGCGCTGCTGAGCTCCACCGCGCGCTGTCGACACCGTC 1314

RESULT 7
AF433156 LOCUS
DEFINITION AF433156 5856 bp DNA linear PLN 23-AUG-2002
Chlamydomonas reinhardtii granule-bound starch synthase I (ST2)
gene, complete cds.
ACCESSION AF433156
VERSION AF433156.1 GI:16755882
KEYWORDS
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonads.
REFERENCE 1 (bases 1 to 5856)
Wattebled,F., Buleon,A., Bouchet,B., Rai,J.P., Lienard,L.,
Delvalle,D., Binderup,K., Dauvillee,D., Ball,S. and D'Hulst,C.
Granule-bound starch synthase I: a major enzyme involved in the
biogenesis of B-crystallites in starch granules
Eur. J. Biochem. 269 (15), 3810-3820 (2002)
1215378
2 (bases 1 to 5856)
Wattebled,F., Ball,S.G. and D'Hulst,C.
Direct Submission
Submitted (15-OCT-2001) UMR8576 du CNRS, Laboratoire de Chimie
Biologique, Universite des Sciences et Technologies de Lille,
Villeneuve d'Ascq CEDEX 59655, France
FEATURES
Location/Qualifiers
1..5856
/organism="Chlamydomonas reinhardtii"
/db_xref="taxon:3055"
1343..>4911
/gene="STA2"
Join(1343..1781,2151..2340,2484..2618,2815..2973,
3186..3373,3582..4117,4295..>4911)
/gene="STA2"
/product="granule-bound starch synthase I"
Join(1480..1781,2151..2340,2484..2618,2815..2973,
3186..3373,3582..4117,4295..4911)
/gene="STA2"
/note="similar to the sequence deposited in GenBank
Accession Number AF026420; GBSSI"
/codon_start=1
/product="granule-bound starch synthase I"
/protein_id="AAL28128.1"
/db_xref="GI:16755883"
/translation="MAVASTSRPSSARPPIVNAASFGVKKTKANLLRELARGSKST
SRSAVTGATGATCALDIDVMAAEVAPWSKTLGLDVTGGLPIELVKRGRHVMIAPRY

```

```

DOYADAWDTSVVVDIMGEKRYFHSIKKGVHRWIDHPWFLAKVWMTGSKLYGPRSG
ADYLDNHRKFAFCFAALEARVLPFGPGEDCVFVANDWHSALVPVLKLDYQPKGQF
TKAKSVLAHNATGQRMWEAFKDTKLPOAFDKLAFSDGYAKYVTEATWEDEKRP
PLTKGTYKINWLGGLIAADKLVTSNYATEADAAGVELDVTYIRAKGIEGIVN
GMDLEENPKTDKFLSVPYDONSVYAKAAAEALQAEGLPVDPTAFLEAFIGLIEE
CGVDIILALPKILATPKVQIAILGTGAAAEKLVNAIGTKYGRAGVYKFSAPLA
HMLTDEADADALATVRASEVAGGRYPENWANCISODLSWSKPAOKWEGILIEVVY
PKGLDEADADALATVRASEVAGGRYPENWANCISODLSWSKPAOKWEGILIEVVY
GKGVATAKKEEIKVPVAEKTGPDLPVSYAPNTLLKPVYASVEGNGAARPKVGTIAPA
MGAMRATTPSGSPSAAATPKVTTTKPALPATAPAKPTAGLAGEASTTSTSENGAASN
GNGNGASASKISAAKPLVSAATKSA*
BASE COUNT 1073 a 1816 c 1852 g 1115 t
ORIGIN

```

```

Query Match 45.3%; Score 962.2; DB 8; Length 5856;
Best Local Similarity 86.2%; Pred. No. 5.2e-87;
Matches 1154; Conservative 0; Mismatches 8; Indels 177; Gaps 1;
QY 963 CACCGTCATCGCGCGCAAGGCGCAITGAGGCAITGTGAACGGCATGGACATGAGGAGTG 1022
Db 3570 CTCCTCCGCGCAGCGCGCAAGGCGCAITGAGGCAITGTGAACGGCATGGACATGAGGAGTG 3629
QY 1023 GAACCCCAAGACCGGACAAAGTTCTCTGTGCGCCCTACGACGACAGACAGGCTCTACGCGG 1082
Db 3630 GAACCCCAAGACCGGACAAAGTTCTCTGTGCGCCCTACGACGACAGACAGGCTCTACGCGG 3689
QY 1083 CAAGGCGCGCGCGCAAGGAGGCGCTCGAGGCGGAGCTGGCGCTGTGGACCCCGCACCGC 1142
Db 3690 CAAGGCGCGCGCGCAAGGAGGCGCTCGAGGCGGAGCTGGCGCTGTGGACCCCGCACCGC 3749
QY 1143 CCGCGCTGTGCGCTTTCATCGCGCGCTGAGGAGGAGGAGGCTGTGGACATCATCTCTGCG 1202
Db 3750 CCGCGCTGTGCGCTTTCATCGCGCGCTGAGGAGGAGGAGGCTGTGGACATCATCTCTGCG 3809
QY 1203 CGCGCTGCGCAAGATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1262
Db 3810 CGCGCTGCGCAAGATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3869
QY 1263 GCGCGCTTACGAGAGCTGTGAGCGCATCGGACCAAGTACAGGCGCGCGCGCGCGCGCG 1322
Db 3870 GCGCGCTTACGAGAGCTGTGAGCGCATCGGACCAAGTACAGGCGCGCGCGCGCGCG 3929
QY 1323 CGTGTCAAGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1382
Db 3930 CGTGTCAAGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3989
QY 1383 GGTGCGCTCGCGCTTCGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1442
Db 3990 GGTGCGCTCGCGCTTCGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4049
QY 1443 CGTGGCGGTGTAGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1502
Db 4050 CGTGGCGGTGTAGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4109
QY 1503 CTTCCACA----- 1510
Db 4110 CTTCCACAGTGAGCGGCGCTTCGCGTGTGTGCTGTGAACITGTGCCCTGAGTGTGCTT 4169
QY 1511----- 1510
Db 4170 GCGGGGCGCGCTTTCGAGCGCGCTTCGACGCTGCTACCGCTGCTGCGCGCGCGCGCG 4229
QY 1511----- 1510
Db 4230 ACCGGTATCGAAGACACCGCTGGCGCATGTGACGCGCTCTCTTTTCCCGCTGTTCCCTGAT 4289
QY 1511-----TGGGCGCGCTGAACCCCGACAGCTGAGCGCTGACGCGCGCGCGCGCGCGCG 1565
Db 4290 CGCAGTGGGCGCGCTGAACCCCGACAGCTGAGCGAGGCTGACGCGCGCGCGCGCGCGCG 4349
QY 1566 CACCGTGGCGCGCTGCGACGCGAGGTGTTTGGCGCGCGCGCTACCGCGAGATGTTGGCCAA 1625
Db 4350 CACCGTGGCGCGCTGCGACGCGAGGTGTTTGGCGCGCGCGCTACCGCGAGATGTTGGCCAA 4409

```

```

QY 1626 CTCATCAGCAGGACTGCTCTGTCCTCAAGCCGCCAGAACTGGGAGGCGCTGCTGGA 1685
Db 4410 CTCATCAGCAGGACTGCTCTGTCCTCAAGCCGCCAGAACTGGGAGGCGCTGCTGGA 4469
QY 1686 GGAGTGGTGTACGGCAAGGCGCGTGGCCACCGCCCAAGAGGAGAGATCAAGTGGC 1745
Db 4470 GGAGTGGTGTACGGCAAGGCGCGTGGCCACCGCCCAAGAGGAGAGATCAAGTGGC 4529
QY 1746 CGTTGGCGAGAAATCCCGGGGACCTGCGCGCGGTGCTCTAGCCCCCAACACCTGAA 1805
Db 4530 CGTTGGCGAGAAATCCCGGGGACCTGCGCGCGGTGCTCTAGCCCCCAACACCTGAA 4589
QY 1806 GCCCGTGTCCCGCTCGGTGAGGCAAGCGCGCGCGCCCAAGGTGCGGCACCAACCGC 1865
Db 4590 GCCCGTGTCCCGCTCGGTGAGGCAAGCGCGCGCGCCCAAGGTGCGGCACCAACCGC 4649
QY 1866 CCCCAGCATGGGCGCGTGGCGCGGACACACCGCTCGGGCGCGCTGCGCGCGCGCCAC 1925
Db 4650 CCCCAGCATGGGCGCGTGGCGCGGACACACCGCTCGGGCGCGCTGCGCGCGCGCCAC 4709
QY 1926 CCCCAGGTGACCACTACAGCCCGCGCTGCGCGCGCCCAAGCGCCAGACCGCTGG 1985
Db 4710 CCCCAGGTGACCACTACAGCCCGCGCTGCGCGCGCCCAAGCGCCAGACCGCTGG 4769
QY 1986 CTTCAAGTGGCGGTGAGGCTCCACCACTCGACCTCGAGAACCGCGCTGCTCCAA 2045
Db 4770 CTTCAAGTGGCGGTGAGGCTCCACCACTCGACCTCGAGAACCGCGCTGCTCCAA 4829
QY 2046 CGGCAAGCAAGCGTGCCTCGGCTCCAGACCTCGGCTGCGCAAGCCCGCTGCTCGC 2105
Db 4830 CGGCAAGCAAGCGTGCCTCGGCTCCAGACCTCGGCTGCGCAAGCCCGCTGCTCGC 4889
QY 2106 CGCACCCGCAAGTCCGCC 2124
Db 4890 CGCACCCGCAAGTCCGCC 4908

RESULT 8
AF286320 2289 bp DNA linear PLN 30-OCT-2000
LOCUS Triticum aestivum granule bound starch synthase I (gbss1) gene,
DEFINITION complete cds.
ACCESSION AF286320
VERSION AF286320.1 GI:11037535
KEYWORDS Triticum aestivum.
SOURCE Triticum aestivum.
ORGANISM Triticum aestivum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
McCue,K.F., Harkman,W.J., Tanaka,C.K. and Anderson,O.D.
1 (bases 1 to 2289)
Starch Branching Enzymes Sbel and Sbe2 from Wheat (Triticum
aestivum cv. Cheyenne): Molecular Characterization, Developmental
Expression, and Homolog Assignment by Differential PCR
Unpublished
2 (bases 1 to 2289)
McCue,K.F. and Anderson,O.D.
Direct Submission
Submitted (11-JUL-2000) United States Department of Agriculture,
Agricultural Research Service, 800 Buchanan Street, Albany, CA
94710-1105, USA
FEATURES
Source
1..2289
/organism="Triticum aestivum"
/cultivar="Cheyenne"
/db_xref="taxon:4565"
/tissue_type="endosperm"
204..2021
/gene="Gbss1"
204..2021
/gene="Gbss1"

```

```

/EC_number="2.4.1.21"
/notes="glucosyltransferase; ADP glucose glucosyl
transferase; waxy; Wx"
/codon_start=1
/product="granule bound starch synthase I"
/protein_id="AAG27624.1"
/db_xref="GI:11037536"
/translation="MAALVTISLATSGTGLITDRFRAGFQVRRSPADAPLMRT
TGASAAPKQSRKRAHRTRCRLSMVVRATSGAGNLIYFVGAEMAPSKYKGGDLVGG
LPMAAANGHRVMIISPRYDQKADWDTSVSEIKVADEYRVRVYFCHYKRGVDFV
DHPCEFLKVRGKTKEITGPDADCTDYEDNOLRESLLQQAALAPRLDLDNNPYSFGP
YGEDVYFVNCNDHRTGLLACYLKSNSGGLYRTAKVAFCHINISYOGRSFDDFAOLN
LPDRFKSFFDIDGDKPVEGRKLNWAGILQADKVLVSPYIAELLSGEARGCEL
DNTMLRTGITGVNGMDVSEWDPAKFLAANDIVTITALEGKALNEALQAEVGLPVD
RKVPLVAFITGRLEEQKPDVMAAIPKLEEDVQIVLLGTGKKRFEERLKSVEEKFP
SKRVAVRFNAPLAHOMAGADVLVTSRPEGLIQLOSMRYGTCPACASTGGLVDT
IMEGKTGFHMGHLSDVNCVVVPEADVKVVTTLKRAVKVGTTPAYHEMVKNMCTQDLISW
KGPAAKRWEDVLLLELGVGSEPGVIGIEIAPLAMENVAAP"
BASE COUNT 480 a 684 c 735 g 390 t
ORIGIN
Query Match 25.3%; Score 537; DB 8; Length 2289;
Best Local Similarity 61.8%; Pred. No. 8.4e-45;
Matches 1011; Conservative 0; Mismatches 530; Indels 96; Gaps 6;
QY 171 CGCGCTGGACATCGTGTGCTGTGAGTGCCTTGGTCCCAAGACGGCGGCT 230
Db 431 CGGCATGAACCTCGTTCGTGGCGCGAGATGCGCCCTGGAGCAAGACGGCGGCT 490
QY 231 GGGCGATGTACTGTGCTGCTGCTATTTAGCTGTGCTCAAGCGCCACCGCTCATGAC 290
Db 491 CGGCGAGCTCTCGGGGGCCCTCCGCCAGCATGCGCCGCAACGGTCAACGGTCACTGGT 550
QY 291 CATTCGCCCTCGCTACGACCACTAGCTGACGCTCGCTGGACACCTCGGTGGTGACAT 350
Db 551 CATCTCCCGCGCTACGACCACTAGCAAGCGCTGGACACCAAGCGCTGCTCGAGAT 610
QY 351 CATGG-----CGGAGAAGTCCGCTACTTCCATCCATCAAGAGGGCGTGA 398
Db 611 CAAGGTCCGCGACGAGTACGAGAGGGTGAAGTACTTCCACTGCTCAACAGCGGGGTGA 670
QY 399 CGGCTGTGGATTGACACCCCTGCTTCTGCGCAAGTCTGGGCAAGACCGGCTCCA 458
Db 671 CGCGTGTCTGTCGACACCGCTGCTTCTTGAGAGAGTCCGGGCAAGACCAAGGAA 730
QY 459 GCTGTACGGCCCCCGCTCCGCGCTGACTACTGTGACACCAACCAAGCGCTTGGCCCTGT 518
Db 731 GATCTACGGCGCGATGCGGCGACGAGTACGAGCAACCAAGCTTACGCTTACGCTGCT 790
QY 519 CTGCNAGCGCGCTATTGAGGCTGCGCGTGTGCTGCTC----- 558
Db 791 CTGCCAGCGCGCTTGAAGGCGCGCGCTGCTGCTGCTC----- 850
QY 559 ---GGCGCGCGGAGGACTCGTCTGCTGGCCACGACTGGCACTCCGCCCTCGGCTGCC 614
Db 851 CGGACCTTACGGGGAAGCGTGTGCTGCTGTCACGACTGGCACAGGCGCTTCTGCG 910
QY 615 CGTCTGTGTGAAGGAGAGTACCGCCCAAGGCGCGCTTACCAAGGCGCAAGTCCGCTGT 674
Db 911 CTGCTACCTCAAGAGCAACTACCACTCCAGTGGCATCTATAGGACGCGCAAGGTAGCGTT 970
QY 675 GCATATCCACACATCGCTTCCAGGCGCGCATGTGGGAGGAGCTTCAAGGACACGAA 734
Db 971 CTGCATCCACACATCTCGTATFAGGGCGCGCTTCTCTTCGAGCACTTCCGCGAGTCAA 1030
QY 735 GGTGCGCGCGCGCTTTTGACAAGCTGTGCGGACGCTATGCCAAGGTATTACAC 794
Db 1031 CCTGCCGACAGGTTCAAGTCTGCTCTGACTTCATCGACGCTAGGACAAG----- 1082
QY 795 TGAGGCCACCCCATGGAGGAGGACGAGAACGCCCGCTGACGGGAAGACCTACAGNA 854
Db 1083 -----CGGTTGGAGGGCGC-----AA 1099

```


QY	855	GATCAACTGGCTGAAGGTGGCAATATATGCGCGCGACAAAGCTGGTGAATGTGTGCGCCAA	914		
Db	1100	GATCAACTGGATGAAGCCGGGATCTCTGACGCCGCAAGAGTGTCTACGGTGAAGCCCTA	1159		
QY	915	CTACGGACAGAGATCGCTGCCGATGCGCGCGGGTGTGGAGCTGGACACCGCTCATCCG	974		
Db	1160	CTACGGGAGAGCTCATCTCCGCGGAAGCAGCGGCTGCGAGCTGCACACATCATCGG	1219		
QY	975	CGCCAAAGGCAATTGAGGGCATGTGAACGGCATGGACATGAGGAGTGAACCCCAAGAC	1034		
Db	1220	CCTCAGGGCATCACCGGCATCTCAACGGCATGGAGCTCAGCGAGTGGAGCCCGCCAA	1279		
QY	1035	CGACAAGTTCCTGTCTGCGCCCTACGACACAGACAGCTCTACCGCGCAAGCGCGCG	1094		
Db	1280	GGACAAGTTCCTGCCGCCAACTACGACGTCAACCCGCGTGGAGGGAAGCGCTGAA	1339		
QY	1095	CAAGSAGGCCCTGAGCGCGAGCTGGCGCTCTGTGGACCCACCGCCGCCCTGTTCGC	1154		
Db	1340	CAAGGAGCGCTGACGCGCGAGGTGGGCTGCGCGTGGACCGGAAGTGCCTGTGGTGC	1399		
QY	1155	CTTCATCGCGCGCTGGAGGACAGAAAGGTGTGGACATCATCTGCGCGCCCTGCCCAA	1214		
Db	1400	CTTCATCGCGAGCTGGAGGACAGAAAGGCCCGCGAGTGTATGATCGCCGCTACCGGA	1459		
QY	1215	GATCCTGGCCACCCCAAGGTGCAGATCGGCATCTCTGGGTACCGGCAAGCGCGCTACGA	1274		
Db	1460	GATCTTGAAGGAGGAGGACCTCAGATCTGTCTCTCGGCGACCGGGAAGAAGTTGA	1519		
QY	1275	GAACTGGTGAAGCCATCGSCACCAAGTACAAGGCGCGCGCAAGGCGGTGTGTCAAGTT	1334		
Db	1520	GCGGCTGCTCAAGAGCTGGAGGAGAGTTCGCGAGCAAGGTGAGGCGCGTGTGTCAAGTT	1579		
QY	1335	CTCGGCGCCCTGGCGCACATGCTCACCGCGCGCGGACTTCACTGCTGTGCGCTCGCG	1394		
Db	1580	CAACGCGCGCTGGCTCACCAGATGATGCGCGCGCGAGTGTCTCCCGTCAACCGCG	1639		
QY	1395	CTTCGAGCCCTGCGGCTGATCCAGCTGCAGCGCATGCGCATACGGTACCGTGCCTGTGT	1454		
Db	1640	CTTCGAGCCCTGCGGCTCATCCAGCTCCAGGGGATGCGCTACGGAACGCGGTGCGCGT	1699		
QY	1455	AGCTCCACCGCGGCTGTGTCACACCGTCAAGGAGGCGGTCAACCGCTTCCACATGGG	1514		
Db	1700	CGGTCCACCGCGGCTGTGTCACACGATCATGGAGGCAACCGGGTTCACATGGG	1759		
QY	1515	CGCCCTGA-----ACCCGACAGCTGGACGAGGTGACCGGACCGCCCTGCGCGCAC	1568		
Db	1760	CCACCTCAGCGTGCATGCAACGTTGTGGAGCGCGCGCGAGTGAAGAAGTGTGTGACCA	1819		
QY	1569	CGTGCCTGTGCGACGAGTGTGTGCGCGCGCGCTACCCGAGATGTGTGCGCAACTG	1628		
Db	1820	CCTGAAGCGCGCTCAAGGTGTGCGCACGCGCAGCTTACCATGATGTGTCAAGACTG	1879		
QY	1629	CATCAGCAGGAGCTGTCTCTGTCCTCAAGCGCGCGCGCAAGTGGAGGCGCTGTGAGGA	1688		
Db	1880	CATGATCCAGATCTCTCTGGAAGGCGCGCGCAAGAACTGGAGGAGCTGTCTTGA	1939		
QY	1689	GTTGGTGTAGCGCAAGGCGCGGTGCGCACCGCCACAGAAAGGAGGAGATCAAGTGCCTG	1748		
Db	1940	ACTGGG---GTCAGGGGAGCGAGCGAGGGTCAATCGGCGAGGAGATTGCGCGCTCGC	1996		
QY	1749	TGCCGAGAGATCCCG 1765			
Db	1997	CATGGAACGTCGCCG 2013			
RESULT 9	AF486514	1812 bp	mRNA	linear	PLN 02-JUL-2002
LOCUS	Hordeum vulgare cultivar oderbrucker granule bound starch synthase				
DEFINITION	I mRNA, complete cds.				
ACCESSION	AF486514				
VERSION	AF486514.1	GI:21667427			
KEYWORDS	.				
SOURCE	Hordeum vulgare.				
ORGANISM	Hordeum vulgare				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Hordeum.				
AUTHORS	1 (bases 1 to 1812) Patron, N., Smith, A., Fahy, B., Hylton, C., Naldrett, M., Rossnagel, B. and Denyer, K.				
TITLE	A mutation in the 5' non-coding region of the barley GBSSI gene alters its temporal and spatial expression and reduces GBSSI activity and amylose content in the endosperm				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1812) Patron, N.				
AUTHORS	Direct Submission				
TITLE	Submitted (22-FEB-2002) Metabolic Biology, John Innes Centre, Norwich Science Park, Norwich, Norfolk NR4 7UH, UK				
JOURNAL	Location/Qualifiers				
FEATURES	1..1812				
source	/organism="Hordeum vulgare" /cultivar="oderbrucker" /db_xref="taxon:4513"				
CDS	1..1812 /note="alpha glycosyl transferase" /codon_start=1 /product="granule bound starch synthase I" /protein_id="AAW74048.1" /db_xref="GI:21667428"				
translating	/translation="MAALATSLQATSGTGLVGTDFRRPFGQLRPNADALGMRT IGASAPKSRKARHGRKSLSVVSATSGMNLVFGAEMAFPHWCTGGIGDVLGLP PMAAAGHRVMVSPRYDQYKDAWDTSVISEKVADEYERVFHCYKRGVDFVFDH PFLKVRGKTEKIYDPAGTDYEDNQRFSLCQAALAPRILNLNNPFSGPYG EDVFFVNCNDWHTGLLACYLSKNYGRFARVAFVHNIISYGFSEFLLISGEARGCELDN DRFKSFDFIDGVDKVEGRKINWKGALIOADKVLTVSPYAEELISGEARGCELDN IMRLTGITGVNMDVSWDPTKDPLAVNVDITTALEAKALKEALQAEVGLPYDRK VPLVAFGLREKQKPDVMIAPLPELKEEDVQILILGIGTKKFKELLKSMEEKFPGR VRVVRFNAPLAHOMMAGADLLAVTSRFPGLIQGMRYGTGTCVCASTGGVDTIV EKGTFHMGRLSDVNCNVPADVKVATTLTKRAVKVYGTGTPAYQEMVKNMIODLSWKG PARNWEDVLELGEVGEPIGVEETAPLAMENVAAP"				
BASE COUNT	363 a 564 c 597 g 288 t				
ORIGIN					
Query Match	25.0%; Score 531.2; DB 8; Length 1812;				
Best Local Similarity	61.6%; Pred. No. 3.4e-44;				
Matches	988; Conservative 0; Mismatches 523; Indels 93; Gaps 5;				
QY	170	GCGCGCTGGACATCGTATGTTGCTGAGTGGCCCTTGTCTCAAGCGCGGCGGCC	229		
Db	221	GCGGATGACCTCTGTTCTGCGCGCGAGATGGCCCTGGAGCAGACCGCGGCC	280		
QY	230	TGGGCGATGTGACTGTGTGGCTGCTTATTGAGTGTGAAGCGCGGCCACCGGTCATGA	289		
Db	281	TTGCGATGTCCTCGCGGCTCCCGCGCCATGGCGGCAACGGTCAACGGGTGATGG	340		
QY	290	CAATTCGCCCTGCTAGGACCATGCTACCGCTGAGCGCTGGACACCTCGGTGGTGACA	349		
Db	341	TGCTCTCCCGCGCTACGATCAGTACAAGGACGCTGGGACCATCCAGCTCATCTCGAGA	400		
QY	350	TCATGG-----GCGAGAAGTCCGCTACTTCCACTCCATCAAGAGGGCGTGC	397		
Db	401	TCAGGTGCTGACGAGTAGGAGGTTGAGGTCTTCCACTGCTACAGCGCGGAGTGG	460		
QY	398	ACCGCTGTGGATTGACACCCCTGTTCTCTGGCCAAAGTCTGGGCAAGACCGGCTCCA	457		
Db	461	ACCGCTGTTCATCGACACCCCTGTTCTCTGGAGAAGTCCGGGCAAGACCAAGGAGA	520		
QY	458	AGCTGTACGCGCCCTCGCGGCTGACTACTGACACCAACCAAGCGCTTCGCCCTGT	517		
Db	521	AGATCTACGCGCGCGCGGACACAGACTATGAGGACCAACCAAGCAGCGCTTCAGCCTTC	580		
QY	518	TCTGCAAGCGCGCTATTGAGGTGCGCGGCTGCTGCCCTTC-----	558		
Db	581	TCGCGCAGGACCATCGAGGACCCAGGATCCTCAACCTCAACAAACACCCCTACTTTT	640		

```
QY 559 -----GGCCCCGGGAGGACTGGTCTTCTGGCCAAAGCTGGCACTCGGCCCTGGTGC 613
Db 641 CTGGTCCCTACGGGAGAGCTGGTCTTCTGGTGAACAGCTGGCACAGGGCCCTCTGG 700
QY 614 CGTCTCGTGTGAAGAGCAGTACCAAGCCCAAGGGCCAGTTCACCAAGGCCAAGTCGGTGC 673
Db 701 CTTGTACCTCAAGAGCAACTACCAAGTCAATGGCATCTACAGAGCGGCCAAGGTGGCT 760
QY 674 TGGCTATCCACACATCGCTTCCAGGGCGCGATGTGGGAGGAGCTTCAAGGACAGA 733
Db 761 TCTGCATCCACACATCTCGTACAGGGCGCTTCTCTCGACGACTTTCGCGACGCTCA 820
QY 734 AGTGGCCCCAGGCGCTTTTGCAAGCTGGCTTCTCGGACGGCTATGCCAAGGTTTACA 793
Db 821 ACCTGCCCGACAGGTTCAAGTCTGCTCTCGACTTCATTACGGCTACGACAAG----- 873
QY 794 CTGAGGCCACCCCATGGAGGAGGAGCAAGAGCCCGCTGACGGGAAGACCTACAAGA 853
Db 874 -----CCGTTGAGGGCGC-----A 889
QY 854 AGATCAACTGGCTGAAGGTTGGCATTATCGCCGCGACAGCTGTGACTGTGTCCGCCA 913
Db 890 AGATCAACTGGATGAAGCGCGGATCTCGAGCCCGACAGGTGCTGACGGTAGCCCT 949
QY 914 ACTACCGCACCGAGATCGCTGCGGATCGCGCGCGGTGTGGAGCTGGACACCGTCATCC 973
Db 950 ACTACGCTGAGGAGCTCATCTCTGGCGAAGCCAGGGCTGCGAGCTCGACACATCATGC 1009
QY 974 GCGCCAAAGGCGATGAGGCGATGTGAACGGCATGTGACATGAGGAGTGAACCCCAAGA 1033
Db 1010 GCGTCAAGGGGATCACCGGAATGTCACAGGCAATGACGCTGAGTGGGACCCACCA 1069
QY 1034 CGACAAAGTTCCTGTGCGCCCTACGACCAAGACAGCTGTACGCGCGCAAGCGCGCG 1093
Db 1070 AGGACAAGTTCCTGCGCGTCACTAGCATACACACCGGTTGGAGGCCAAGCGCTGA 1129
QY 1094 CCAGGAGGCGCTTGCAGGCGAGCTGGGCTGCTGTGGACCCCAAGCGCCCTCTGTCG 1153
Db 1130 ACAAGAGGCGCTGACGCGCGAGTGGGCTGCGGCTGGACAGGAGTGGCGCTGGTG 1189
QY 1154 CTTTCATCGCCCGCTGGAGGAGCAAGAGGTGTGACATCATCTCTGGCGCGCTGCCCA 1213
Db 1190 CTTTCATCGCAGCTGGAGGAGCAAGAGGCGCCGACGCTGATGTCGCGCGCATCCGG 1249
QY 1214 AGATCTGCGCACCCCAAGGTCAGATCGCATCTCTGCGTACCGGCAAGCGCGCTACG 1273
Db 1250 AGATCTGAGGAGGAGGCTCCAGATCATCTCTCTGCGCACCGGGAAGAAGTTTG 1309
QY 1274 AGAAGCTGTTGAAGCGCATGGCACCAAGTACAGAGGCGCGCCCAAGGCGGTGGTCAAGT 1333
Db 1310 AGAAGCTGCTCAAGAGCATGGAGGAGAAGTTCCCGGCAAGGTGAGGCGCGTGGTCAAGT 1369
QY 1334 TCTCGGCGCCCTGGCGCATCTCTACCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTG 1393
Db 1370 TCAAGCGCGCTGAGTCTACAGATGATGGCGCGCGCGGCTGCTGCTGCTGCTGCTG 1429
QY 1394 GCTTCGAGCGCTGCGGCTGATCCAGCTGACGCCATGACGCTACGCTACGCTGCGCGCTG 1453
Db 1430 GCTTCGAGCGCTGCGGCTCATCCAGCTCCAGGGAATGCGCTATGNAACCGCTGGGTG 1489
QY 1454 TAGCTTCCACCGCGCGCTGGTTCGACACCGCTCAAGGAGGCGGCTACCGGCTTCCACATGG 1513
Db 1490 GCGGCTCCACCGCGCGCTGCTCGACACGATGCTGGAGGCAAGACCGGCTTCCACATGG 1549
QY 1514 GCGCCCTGA-----ACCCGCAAGCTGACGAGGCTGACCGCGAGCCCTGGCGCCCA 1567
Db 1550 GCGGCTTACGCTGCAAGCTGTTGGAGCGCGGAGCTGGAAGAGTGGCGGACCA 1609
QY 1568 CGTGCGCGCTGCGACGAGGTTTTCGCGCGCGCGCTACCCCGGAGATGGTGGCGCAACT 1627
Db 1610 CCTGAAGCGGCGCTCAAGGTCGTGCGGCACCGCGCGTACCAGGAGATGGTCAAGAACT 1669
```

```
QY 1628 GCATAGCCAGGACCTGTCTCTGTTCAAGCCGCCAGAGTGGGAGGCGCTCTCTGAGG 1687
Db 1670 GCATATCCAGGATCTCTCTGGAAGGAGCTGCCAAGAACTGGGAGGACGTGCTTCTGG 1729
QY 1688 AGTGTGTGTACGCAAGGCGGTGGCCACCCGCCAAGAAGGAG 1731
Db 1730 AACTGGGGTGGAGGAGCGAGCGGGGATCGTCGCGGAGGAG 1773

RESULT 10
HWAXYR
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
FEATURES
source
/organism="Hordeum vulgare subsp. vulgare"
/strain="vogelsanger Gold"
/db_xref="taxon:112509"
/map="chromosome 1"
/clone="pcwx27"
/clone_lib="EMBL4"
226..2037
/note="starch synthase"
/codon_start=1
/protein_id="CAA30756.1"
/db_xref="GI:19129"
/db_xref="SWISS-PROT:P09842"
/translacion="MAALATSLATSGTLVGTDRFRPFGFQGLRPRNPADALGMRT
IGASAAPKQSRKAHGRSRLSVVSATSGSMNLVFGAEMAPWSKTIKGLDVLGLP
PAMAANGHRVMSVSPRYDQKDAWDTSVSEIKVADEYERVRFFHCYKRGVDFIDH
PWFLKVRGKTKEIKYDPADGTDYEDNOQRFSLCQAALAPRILNINNNPFSGYPG
EDVFCNDWHTKLLACYLKSNTSGNIVRTAKVAFCHINISYQGRFSDFDFAQLNLP
DRPKSDFDIDGDKPEGRKLNWKGILQADKVLTVSPYAFELISGEARGCLDN
IMRLGITGVNGMVSNDPTKDFLAVNYDITLLEAKALKEALQEVGLPVDNR
VPLVAFGLRLEKQKDPVMAIPILKEEDVOILLGDKKAFKLLKMSMEKFFPK
VRVAFNAPLAHOMMAGADLLAVTSRFPFQGLIQGMRYGTPCVCASTGGGLDFTIV
EGKTFGHMGRLSVDCNVVPPADVKVATTKIKRAVKVYVTPAYQEMVKMCMQIDLSWK
PAKNWEDVLELGVGSEPIGVEIAPLAVMENVAAP"
2276
/note="poly A site (pcwx4)"
2283
/note="poly A site (pcwx27)"
2301
/note="poly A site (pcwx18)"
BASE COUNT 501 a 688 c 714 g 403 t
ORIGIN
Query Match 25.0%; Score 531.2; DB 8; Length 2311;
Best Local Similarity 61.6%; Pred. No. 3.1e-44;
Matches 988; Conservative 0; Mismatches 523; Indels 93; Gaps 5;
QY 170 GCGCGTGGACATCGTATGTTGCTGCTGAGGTGCGCCCTTGGTCCAAAGCGGCGGCC 229
```

Db 446 GCGCATGAACCTCGTGTTCGGCGCCGAGATGGCGCCCTGAGCAAGACCGCGGCC 505
QY 230 TGGCGATGTGACTGGTGGCGTGGCTATTGAGCTGGTCAAGCGCGGCCACCGCGCTCATGA 289
Db 506 TTGCGATGCTTGGCGGCCCTCCCGCGCCATGGCGCCCAACGGTCAACCGGTGATGG 565
QY 290 CCATTGCGCCCTGATGACACAGTACGCTGACGCTGGGACACCTCGGTGGTCTGTGACA 349
Db 566 TCGTCTCCCGCGCTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 565
QY 350 TCATGG-----GGGAGAGTCCGCTACTTCCACTCCATCAAGAGGCGGTGC 397
Db 626 TCAGGTGCTGACGAGTACGAGAGGTGAGGTTCCTCCACTGCTCAAGCGCGGATGG 685
QY 398 ACCCGGTGTGATGACACCCCTGGTTCCTGGCAAGGTGTGGGCAAGACCGGCTCCA 457
Db 686 ACCCGGTGTTCATGACACCCGCTGGTTCCTGGCAAGGTTCGGGCAAGACCAAGGAGA 745
QY 458 AGCTGTAGGCGCCCGCTCGGCGCTGACTACCTGGACACCAAGCGGTTCGCGCTGT 517
Db 746 AGATCTACGGCGCGCGCGCGGACAGACTATGAGGACCAACGACGCGCTTCAGCGCTC 805
QY 518 TCTCAAGGCGCTATTGAGGCTGCCCGGCTGCTGCGCTTC----- 558
Db 806 TCTGCCAGGACGACTCGAGGCCACCGAGTCTCACTCACTCAACCAACACCCCTACTTT 865
QY 559 -----GGCGCGGAGGAGTGGTGTTCGTGGCAAGCTGACGACTCGCGCTCGGTGC 613
Db 866 CTGCTCCTACGGGGAAGAGTGGTGTTCGTGCAACGACTGACACGCGCTTCGTG 925
QY 614 CCGTCTGCTGAGGAGGAGTACGAGCCCAAGGCGAGTTCACCAAGGCGCAAGTCGTTGC 673
Db 926 CCGTCTACCTCAAGAGCAATACAGTCCAAATGGCATCTACAGGACGCGCAAGGTGGCT 985
QY 674 TGGTATATCCCAACATCGCTTCAGGCGCGCATGTGGGAGGCGCTTCAAGGACACGA 733
Db 986 TCTGATCCACACATCTGTTACGAGGCGGCTTCTCTCTCGAGAGCTTTCGCGAGCTCA 1045
QY 734 AGCTGCGCCGAGCGCGCTTTGACAAAGTGGCTTCTCGAGCGGTATGCCAAGGTTTACA 793
Db 1046 ACCTGCGCGAGAGTTCAAAGTCTGCTTCGACTTCATTGAGGCTACGACAAG----- 1098
QY 794 CTGAGGCCACCCATGAGGAGGACGAGAACGCGCGCTGACGGGAAGACCTTACAAGA 853
Db 1099 -----CCGTGGAGGGGCGC-----A 1114
QY 854 AGATCAACTGGCTGAAGGGTGGCATATCGCGCGGACAGCTGGTGAATGTGTGCGCCA 913
Db 1115 AGATCAACTGGATGAAGCGCGGATCTCGAGGCGGACAGAGGTGCTGACGGTGAAGCCCT 1174
QY 914 ACTACGGACCGGATCGCTGCGGATGCGCGCGCGGTGTGGAGCTGGACACCGCTCATCC 973
Db 1175 ACTACGCTGAGGAGCTCATCTGCGCAAGCCAGGGGCTGCGAGCTCGACAACATATGC 1234
QY 974 GCGCAAGGCGATTGAGGCGATTGTAAGCGCATGGACATGAGGAGTGAACCCCAAGA 1033
Db 1235 GCCTACGGGGATCACCGGAATGCTCAACGCGATGGAGCTGAGTGGGAGCCCGCCACA 1294
QY 1034 CCACAAAGTTCCTGTCGCGCTTACCAACCAAGACAGCGTCTACGCGCGCAAGCGCGCG 1093
Db 1295 AGGACAAGTTCCTGCGCGCTCACTACGACATCACCGCGGTGGAGCGCAAGCGCTGA 1354
QY 1094 CCAAGGAGCGCTCGAGCGGAGTGGGCTGGCTGTGGACCCACCGCGCGCTGTGTCG 1153
Db 1355 ARAAGGAGCGCTCGAGCGGAGTGGGCTGCGGCTGGACAGAGGTGCGCGTGTGG 1414
QY 1154 CTTTCATGCGCGCTCGAGGAGCAGAGGAGTGTGGACATCATCTCGCGCGCTGCGCA 1213
Db 1415 CTTTCATGCGGAGCTCGAGGAGCAGAGGAGCGCGCGAGTGTGATGATCGCGCGCTACCGG 1474
QY 1214 AGATCTTGGCCACCCCAAGGTGAGATCGCATCTCTGGGTACCGGCAAGCGCGCTACG 1273

Db 1475 AGATCTGAAGGAGGAGGACGTCAGATCATCTCTTGGCACCGGGAAGAAGTTTG 1534
QY 1274 AGAAGCTGTGAAGCCCATCGCACCAAGTACAGGCGCGCGCCAGCGTGGTCAAGT 1333
Db 1535 AGAAGCTGTCAAGAGCATGAGGAGAAAGTTCGCGGCAAGTGAAGCGCGTGGTCAAGT 1594
QY 1334 TCTCGGCGCCCTGGCGACATGCTACCGCGCGCGGCACTTCACTGCTGGTGGCTCGC 1393
Db 1595 TCAAGCGCGCTAGCTCAACAGATGATGCGCGCGCGGCACTTGTGCTGCTGCAACGCG 1654
QY 1394 GCTTCGAGCCCTGCGGCTGATCAGCTGACGCGCATGACACTACGCTACGCTGCGCGTGG 1453
Db 1655 GCTTCGAGCCCTGCGGCTCATCAGCTCCAGGAATCGCTATGTAACGCGCTGCTGT 1714
QY 1454 TAGCTTCACCGCGCGCTGCTGACACCGTCAAGAGGCGGCTACCGGCTTCCACATGG 1513
Db 1715 GCGGCTCCACCGCGCGCTGCTGACAGATCTGAGGAGGCAAGACCGGCTTCCACATGG 1774
QY 1514 GCGCGCTGA-----ACCGCGACAGCTGAGGAGTGAAGCGCGGCGCGCGCGCA 1567
Db 1775 GCGGCTCAGCGTGCAGCTGCAACGCTGAGCGCGCGGAGCTGAAGAGGCTGCGGACCA 1834
QY 1568 CCGTGGCGCGTGCAGGAGGCTGTTGGCGGCGCGCTACCGGAGTGTGGCGCACT 1627
Db 1835 CCCTGAAGCGCGCGCTCAAGGTCTGCGCACGCGCGGCTACAGGAGATGCTCAAGAACT 1894
QY 1628 GCATCAGCGAGGAGCTGCTCTGCTGCTCAAGCGCGCGGAGTGGAGGCGCTGCTGGAGG 1687
Db 1895 GCATGATCCAGGATCTCTCTGGAAGGAGCTGCAAGAACTGGGAGGACGCTGCTCTGG 1954
QY 1688 AGGTGGTGTACGCGAAGGCGCGGTGGCGCGCGCGCGCGCGCAAGAGGAG 1731
Db 1955 AACTGGGCGTGGAGGAGCGAGCGCGGATCGTGGCGGAGGAG 1998

RESULT 11

AF486515

LOCUS

DEFINITION

AF486515

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

1..1827

/organism="Hordeum vulgare"

/cultivar="waxy oderbrucker"

/db_xref="taxon:4513"

1..1827

/note="alpha glycosyl transferase"

/codon_start=1

/product="granule bound starch synthase I"

/protein_id="AA074049.1"

/db_xref="GI:21667430"

/translation="MAALATSLATSGTGLVTDTRAPSMFRAGTQGLRPRKPADGT

FGMRTVGASAAPKQSRKAHRNRCLSVVVRATGSGMNLVFGAEMAPWSKTGLGVD

1827 bp mRNA linear PLN 02-JUL-2002

Hordeum vulgare cultivar waxy oderbrucker granule bound starch

synthase I mRNA, complete cds.

AF486515

AF486515.1 GI:21667429

Hordeum vulgare.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Pooideae; Triticeae; Hordeum.

1 (bases 1 to 1827)

Patron, N., Smith, A., Fahy, B., Hylton, C., Naldrett, M., Rossnagel, B.

and Denyer, K.

A mutation in the 5' non-coding region of the barley GBSI gene

alters its temporal and spatial expression and reduces GBSI

activity and amylose content in the endosperm

Unpublished

2 (bases 1 to 1827)

Patron, N.

Direct Submission

Submitted (22-FEB-2002) Metabolic Biology, John Innes Centre,

Norwich Science Park, Norwich, Norfolk NR4 7UH, UK

Location/Qualifiers

1..1827

/organism="Hordeum vulgare"

/cultivar="waxy oderbrucker"

/db_xref="taxon:4513"

1..1827

/note="alpha glycosyl transferase"

/codon_start=1

/product="granule bound starch synthase I"

/protein_id="AA074049.1"

/db_xref="GI:21667430"

/translation="MAALATSLATSGTGLVTDTRAPSMFRAGTQGLRPRKPADGT

FGMRTVGASAAPKQSRKAHRNRCLSVVVRATGSGMNLVFGAEMAPWSKTGLGVD

LGGLPMAAANGHRVNVSPRYDQYKDAWDTSVISEIKVADEYERVRFFHCYKRGVDR
 VPIDEWFLEKRGKTKERIYIGPDAGTQEDNQRFSLCQALEPHNINNNPFF
 SGPYGDVFNQNDHTKGLIYACVSNQNGIYRTAKVAFCLHNSYGRFSFDDFA
 OLNLPREKSSDFIDGDKPVEGRKINMKAGILOADKVLTVSPYAEELISGEARG
 CELNDMLRLTGITGVNGMDVSEWDFTKDFLAYNDITITALEKALKEALQAEVGL
 PVDKPLVAFITGLEQKQPDVMAIAPEILKEEDYQIILITGKKKFKELKLSWEE
 KPFGRVAVRVNAPLAHOMAGADLLVTSRFEPCGLIOLQGMRYGTFCVCASTGGL
 VDTIVEGKTFWHRGLSDVNCNVPEADVRKVTATLKRVRVGTGPAYQBMVKNMCIQD
 LSWGPAKWNEDVLLLELGVESGEPGIVGEIAPLAMENVAAP*
 BASE COUNT 364 a 567 c 602 g 294 t
 ORIGIN

Query Match 24.9%; Score 529.6; DB 8; Length 1827;
 Best Local Similarity 61.5%; Pred. No. 4.8e-44;
 Matches 987; Conservative 0; Mismatches 524; Indels 93; Gaps 5;

QY 170 GCGCGTGGACATCGTATGTTGCTGCTGAGTTCGCCCCCTTGTCTCAAGACGGCGGCC 229
 DB 236 GCGGCATGAACCTCGTCTGTCGGCGCGAGATGGCGCCCTGGAGCAAGACCGCGGCC 295
 QY 230 TGGCGATGTGACTGTGGCTGCTATTGAGCTGGTCAAGCGCGGCACCGCGTCAIGA 289
 DB 296 TTGGGAGCTTCCTGGCGGCTTCCACACGATGGCGGCCAAGCGTACCGGGTGAIG 355
 QY 290 CCATTGCCCTCGCTACGACAGTACGCTGACGCTGGGACACCTCGGTGCTGTGACA 349
 DB 356 TCGTCTCCCGCGCTACGATCAGTACAGGAGCGCTGGGACACCGAGCTCATCTCCGAGA 415
 QY 350 TCATGG-----GGAGAGTTCGCTTACCTCCACTCCATCAAGAGGGCGTGC 397
 DB 416 TCAAGGTTCGCTGAGTACGAGAGGTGAGGTTCTTCCACTGTACAGCGCGGAGTGG 475
 QY 398 ACCGCTGTGATGACACCCCTTGTCTTGGCCAAAGTCTTGGGCAAGACCGGCTCCA 457
 DB 476 ACCGCTGTTCATGACACCCCTGTTCTTGGAGAGTTCGCGGCAAGACCAAGAGA 535
 QY 458 AGCTGTACGGCCCGCTCGGCGCTGACTACCTGGACACCAAGCGTTCGCCCTGT 517
 DB 536 AGATCTACGGCGCGCGCGGCGGACGACTACGAGGACCAACGAGCGGTTACGCGTTC 595
 QY 518 TCTCAAGGCGGCTATTGAGCTGCCGCGTGTGCGCTTC----- 558
 DB 596 TCTGCCAGGAGCGCTCGAGGACCCAGGATCCTCAACCTCAACAAACCCCTACTTTT 655
 QY 559 -----GGCCCGCGGAGGACTCGCTTCTGGCCAAAGACTGGCACTCGCCCTGTGTC 613
 DB 656 CTGCTCCTACGCGGGAAGACGTGTGTGTGTGTCACGACTGGCACACGCGGCTTCTGG 715
 QY 614 CCGTCTCTGAGGACGAGTACAGCCCAAGGCGGAGTTCACCAAGGCCAAGTCGGTGC 673
 DB 716 CCGTCTACCTCAAGCAACTACCAAGTCCATGGCATCTACAGGACGCGCAAGGTGCGCT 775
 QY 674 TGGCTATCCCAACATCGCTTCCAGGCGCGCATGTGGGAGGCGTTTCAAGGACACGA 733
 DB 776 TCTGCATCCCAACATCTGTACAGGCGCGCTTCTCTTCGACGACTTTCGCGAGCTCA 835
 QY 734 AGTGCCTCCAGGCGCGCTTTGACAGTGGCGCTTCTCGAGCGGCTATGCCAAGTTTACA 793
 DB 836 ACCTGCCCGGACAGTTTCAAGTTCGCTTCTCGACTTCAATGACGCGGTACGACAAG----- 888
 QY 794 CTGAGGCCACCCCATGAGGAGGACGAGGAGGCGGCTGACGCGGAAGACCTTACAAGA 853
 DB 889 -----CCGCTGAGGGGCGC-----A 904
 QY 854 AGATCAACTGGCTGAAGGGTGGCATATTTCGCGCGCCACAAAGCTGGTACTGTTCGCCCA 913
 DB 905 AGATCAACTGGATGAAGGCGGGATCTCTGCAGGCGGACAAAGTCTGACGTGAGCCCT 964
 QY 914 ACTACGACGAGATCGCTGCCGATCGCGCGCGGTGTGGAGCTGGACACCGCTCATCC 973
 DB 965 ACTACGCTGAGGAGCTCATCTCTGGCGAAGCGAGGCGCTGCGAGCTCGACAACATCATGC 1024
 QY 974 GCGCAAGGGCATTGAGGGCATTGTGAACGCGCATGGAGGTGGAGGAGTGGAAACCCCAAGA 1033

DB 1025 GCCTACGGGGATCACCGGAATCTCAACGCGATGGAGCTCAGTGGAGCCCA 1084
 QY 1034 CCGACAAGTTCCTGCTCGCGCCCTACGACAGACAGCTCTACGCGGCAAGCGCGCG 1093
 DB 1085 AGGACAAGTTCCTCGCGCTCAACTACGACATCACCCGCGTTGAGGCCAAGGCGCTGA 1144
 QY 1094 CCAAGGAGGCGCTGCAGGCGGAGCTGGGCGCTGCTGTGGACCCACCGCCCTGTTCG 1153
 DB 1145 ARAAGGAGGCACTCGAGGCGGAGTGGGCGTGGCGGTGGACAGGAAGTGGCTGTGG 1204
 QY 1154 CTTTCATGGCGCTCGAGGAGCAGAAGGTGTGGACATCATCTGGCCGCCCTGCCCA 1213
 DB 1205 CTTTCATGGCAGCTCGAGGAGCAGAAGGCGCGGATGATGATGCGGCCCATCCGG 1264
 QY 1214 AGATCTGGGCGCACCCCAAGGTGCAGATCGCATCTCTGGGTACCGGCAAGCGCTACG 1273
 DB 1265 AGATCTTGAAGGAGGAGGAGCTCCAGATCATCTCTTGGCCAGCGGGAAGAAGTTTG 1324
 QY 1274 AGAAGCTGGTGAACCGCATCGGCACCAAGTACAAGGGCGCGGCAAGGCGGTGGTCAAGT 1333
 DB 1325 AGAAGCTGCTCAACAGCATGGAGGAGAAGTTCCCGGGCAAGGTGAGGCGGTGTGAGT 1384
 QY 1334 TCTCGGCGCCCTGGGCGCATGCTCACCGCGCGGCGGAGTTCATGCTGTGCGCTGCGC 1393
 DB 1385 TCAACGCGCGCTAGCTCACGAGATGATGGCGGCGCGGACTTGTCTGCTTCACAGCC 1444
 QY 1394 GCTTCGAGCGCTGGGCGCTGATCCAGCTGCAGCGCATGCTACCTAGCTACCGTACCGTCCG 1453
 DB 1445 GCTTCGAGCGCTGGGCGCTCATCCAGCTTCCAGGGAATTCGCTATGGAAACCGCTGCGTGT 1504
 QY 1454 TAGCTCTACCGCGGCGCTGTGTCGACACCGTCAAGGAGGCGGTCAACGCGTTCACATGG 1513
 DB 1505 GCGCGTCCACCGCGGCTGCTGCGACAGATCGTGGAGGGCAAGACCTGTTCACATGG 1564
 QY 1514 GCGCGCTGA-----ACCCGCAAGCTGGAGCGGCTGACGCGGAGTGGCGGCA 1567
 DB 1565 GCGCGCTCAGCTGCACTGCAACGTGTGGAGCGCGGCGGAGTGAAGAAGTGGCGGCA 1624
 QY 1568 CCGTGGCGCGTGCAGGCGGCTGTTGCGGCGCGGTACCCCGAGATGTTGGCCAACT 1627
 DB 1625 CCCTGAAGCGGCGCTCAAGTCTGCGCACGCGCGCTACGAGAGTGTCAAGACT 1684
 QY 1628 GCATCAGCCAGGAGCTGCTGTCAGAGCGCGCCAGAGTGGAGGCGCTGTGGAG 1687
 DB 1685 GCATGATCCAGGATCTCTCTGGAAGGACCTGCAAGACTGGGAGGAGCTGTCTTGG 1744
 QY 1688 AGTGTGTAGCGCAAGGCGGCTGGCCACCGCCCAAGAGAG 1731
 DB 1745 AACTGGGGGTGGAGGAGCGAGCGCGGATCGTGGCGGAG 1788

RESULT 12
 AF486518
 LOCUS AF486518 1827 bp mRNA linear PLN 02-JUL-2002
 DEFINITION Hordeum vulgare cultivar SB 85750 granule bound starch synthase 1
 mRNA, complete cds.
 ACCESSION AF486518
 VERSION AF486518.1 GI:21667435
 KEYWORDS
 SOURCE Hordeum vulgare.
 ORGANISM Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticeae; Hordeum.
 REFERENCE 1 (bases 1 to 1827)
 Paton, N., Smith, A., Fahy, B., Hylton, C., Naldrett, M., Rossnagel, B.
 and Denyer, K.
 TITLE A mutation in the 5' non-coding region of the barley GBS1 gene
 alters its temporal and spatial expression and reduces GBS1
 activity and amylose content in the endosperm
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1827)

Db	1190	CTTCTCAGGCTGGAGGAGCAAGGCGCCGACGTGATCGCGCCATCCGG	1249
QY	1214	AGATCTGGCCACCCCAAGTGCAGATCCCATCTCTGGTACCGCAAGCCGCTACG	1273
Db	1250	AGATCTCTGAAGGAGGAGGACCTCCAGTATCTCTCTGGCAGCGGAAGAAGTTG	1309
QY	1274	AGAAGCTGGTGAAGGACATCGGCACCAAGTACAAAGGCGCGCAAGGGCGTGTCAAGT	1333
Db	1310	AGAGCTGCTCAGAGCATGAGGAGAGTTCGCGGCAAGGTGAGGGCGTGTCAAGT	1369
QY	1334	TCGCGGCGCCCTGGCGCACATGCTCACCGCGCGGCGCCGACTTCATGCTGGTCCGCTCGC	1393
Db	1370	TCAACGCGCGCTAGCTCAGAGATGATGGCGCGCGCGGACTTGCCTGCTCACCAGCC	1429
QY	1394	GCTTCAGCGCTCGCGCTGATCCAGCTGACCGCATGAGGCGGCGGCGGCGTGTCAAGT	1453
Db	1430	GCTTCAGCGCTCGCGCTGATCCAGCTGACCGCATGAGGCGGCGGCGGCGTGTCAAGT	1489
QY	1454	TAGCCTCCACCGCGCGCTGGTGCACACCGTCAAGGAGGCGGCTACCGGCTTCCACATGG	1513
Db	1490	GCAGCTCCACCGCGCGCTCGTGCACACGATCGTGGAGGCGCAAGACCGGCTTCCACATGG	1549
QY	1514	GCSCCTGA-----ACCCGACAGCTGACGAGGCTGACGCGAGCGCCCTGCGCCGCA	1567
Db	1550	GCGCGCTCAGCTGCACTGCAACGCTGGTGGAGCGCGCGGACGTGAAGAAGGTGCGGACCA	1609
QY	1568	CCGTGCGCGCTGCGAGGAGCTTTCGCGGCGCGCGCTACCCGAGATGCTGCGCAACT	1627
Db	1610	CCCTGAAGCGCGCTCAAGTCTGCGGACGCGCGGCTACAGGAGATGCTCAAGACT	1669
QY	1628	GCATACGCGAGGACCTGCTCTGCTGCAAGCGCGCGCGGAGGAGGCGCTGCTGAGG	1687
Db	1670	GCGTGTCCAGGACTCTCTCGAAGGAGCTGCGCAAGAACTGGGAGGACGTGCTTCTGG	1729
QY	1688	AGTGCTGTACGCAAGGCGCGCTGCGCCACGCGCCAGAGAGGAG	1731
Db	1730	ACTGGGGTGGAGGAGGAGCGCGGGGATCGTGGCGGAGG	1773
RESULT 15			
AF486516			
LOCUS			
DEFINITION			
Hordeum vulgare cultivar Yon M kei non-functional granule bound			
starch synthase I, complete sequence.			
ACCESSION			
AF486516			
VERSION			
AF486516.1 GI:21667431			
KEYWORDS			
SOURCE			
Hordeum vulgare.			
ORGANISM			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
Poideae; Triticeae; Hordeum.			
REFERENCE			
1 (bases 1 to 1827)			
AUTHORS			
Patron, N., Smith, A., Fahy, B., Hylton, C., Naldrett, M., Rossmagel, B.			
and Denyer, K.			
A mutation in the 5' non-coding region of the barley GBSSI gene			
alters its temporal and spatial expression and reduces GBSSI			
activity and amylose content in the endosperm			
Unpublished			
2 (bases 1 to 1827)			
JOURNAL			
Patron, N.			
REFERENCE			
Direct Submission			
AUTHORS			
TITLE			
Submitted (22-FEB-2002) Metabolic Biology, John Innes Centre,			
Norwich Science Park, Norwich, Norfolk NR4 7UH, UK			
JOURNAL			
Location/Qualifiers			
FEATURES			
1..1827			
/organism="Hordeum vulgare"			
/cultivar="Yon M Kei"			
/db_xref="taxon:4513"			
1..1827			
/note="contains non-functional granule bound starch			
synthase I due to premature stop codon"			
misc_feature			

BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity 24.8%; Score 526.4; DB 8; Length 1827;			
Matches 985; Conservative 0; Mismatches 526; Indels 93; Gaps 5;			
QY	170	GCAGCGTGGACATCGTGAIGGTTGCTGCTGAGTGC3CCCTTGTTCACAGACGCGCGGCC	229
Db	236	GCAGCATGACCTGCTGTTCTGTCGGCGCGAGATGCGCCCTGGAGCAAGACCGCGGCC	295
QY	230	TGGCGGATGTAAGTGGTGGCTGCTATGTAAGTGGTTCAAAGCGCGGCAACCGGCTCATGA	289
Db	296	TTGGCGAGCTCTCGCGCGGCTTCCACACGACCATGTCGCGGCAACCGGTCAGCGGTGATGG	355
QY	290	CCATGCGCCCTCGCTAGCAGCAGTACGCTGACGCTGGGACACCTCGTGTCTGTGACA	349
Db	356	TCGTCTCCCCGCGCTAGCATGACAGGACGCTGGGACACCGGCTCATCTCCGAGA	415
QY	350	TCATGG-----GCGAGAAAGTCCGCTACTTCCATCCATCAAGAAGGCGCTGC	397
Db	416	TCAGGTGCTGACGAGTACGAGAGGTGAGGTTCTTCCACTGCTACAAGCGCGAGTGG	475
QY	398	ACCGGTGTGATGACACACCCCTGTTCTGCGCCAAAGTCTGGGGCAAGACCGGCTCCA	457
Db	476	ACCGGTGTGATGACACACCCCTGTTCTGCGCCAAAGTCTGGGGCAAGACCGGCTCCA	535
QY	458	AGCTGTAGCGCCCGCTCGCGGCTGCTACCTGCTGACACCAACCAAGGCTTCCGCTGT	517
Db	536	AGATCTAGCGCCCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	595
QY	518	TCTCAAGGCGCTGTTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	558
Db	596	TCGCCAGGAGCGCTCGAGGACCCAGAGTCTCAACCTCAACCAACCAACCTTCTTT	655
QY	559	-----GGCCCCGGGAGGAGTGGCTCTGCTGCGCCAAAGGCTTCAACCAAGGCTTCA	613
Db	656	CTGTCTCTTACGGGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	715
QY	614	CCGTCTCTGCTGAAGGAGGAGTACCAAGGCTTCAACCAAGGCTTCAACCAAGGCTTCA	673
Db	716	CTGCTACCTCAAGCAACTACCACTCCATGCTACAGGACGCGCAAGGTGGCT	775
QY	674	TGGTATCCAAACATCGCTTCCAGGCGCGCATGTCGAGGAGGCTTCAAGACACGA	733
Db	776	TCTGATCCACACATCTCGTACCAAGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	835
QY	734	AGCTGCCCGCGCGCTTGTGACAAAGTGGCTTCTGCGGAGGCTTCAACCAAGGCTTCA	793
Db	836	ACCTGCCCGAGAGGTTCAAGTCTGCTTCACTTCACTTCACTTCACTTCACTTCACTTCA	888
QY	794	CTGAGGCGACCCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	853
Db	889	-----CCGCTGGAGGGGCGC-----A	904
QY	854	AGATCACTGGCTGAAGGTTGGGCAATATCGCGCGCGCGCAAGCTGGTGGTGGTGGTGG	913
Db	905	AGATCACTGGTGAAGGCGCGGATCTCTGAGGCGCGCAAGGCTGCTACGCTGAGCCCT	964
QY	914	ACTACGCGAGGAGTCTGCTGCGGATCGCTGCGGAGGCTGCTGCGGAGGCTGCTGCGG	973
Db	965	ACTACGCTGAGGAGTCTATCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1024
QY	974	GCGCCAAGGCTTGAAGGCTTGTGAACGCGCATGACATTTAGGAGTGGAAACCCCAAGA	1033
Db	1025	GCCTCACGGGATCACCGGATCGTCAACGCGCATGAGGCTCAGTGTGAGTGGACCCACCA	1084
QY	1034	CCGACAAAGTCTCTGCTGCGGCTTACGACACAGAGGCTTACGCGGCAAGGCGCGCG	1093
Db	1085	AGGACAAAGTCTCTGCGGCTCAACTACGACATCACCAAGGCTTGGAGGCGGAGGCTGA	1144
QY	1094	CCAGGAGGCGCTCAGGCGGAGCTGGGCTTGTGCTGTGAGACCCACCGCGCGCGCTGTG	1153

Db 1145 ACAAGGAGGACATGCGAGCGGAGGTGGGCTGCCGTGGACTGGAAGGTGCCGCTGGTGG 1204
QY 1154 CTTTCATCGGCCCGCTGGAGGAGCAGAAGGGTGTGCATCATCTCTGGCGCGCCTGCCCA 1213
Db 1205 CCTTCATCGGCAGGCTGGAGGAGCAGAAGGGGCCCGCCACCTGATGATCGCGCTCATCCCGG 1264
QY 1214 AGATCCTGGCCACCCCGCCCAAGGTGCAGATGCCATCTGCTGGTACCGCGCAAGGGCGCCCTACG 1273
Db 1265 AGATCCTGAAGGAGGAGGACGTCACAGATCATCTCTGGCACCGGGAAGAAGATTG 1324
QY 1274 AGAAGCTGGTGAACGCCATCGGCACCAAGTACAAAGGGCGCGCCAAAGGGCGTGGTCAAGT 1333
Db 1325 AGAAGCTGTCAAGAGCATGAGGAGAAGTTCCTGGGCAAGGTGAGGGCCGTGGTCAAGT 1384
QY 1334 TCTCGGCGCCCTGGCGGCATGCTCACCGCGGCGCGCCGACTTCATGCTGGTGCCCTCGC 1393
Db 1385 TCAACGCGCGGTAGTCAACAGATGATGGCGCGCGCGACTGCTCGCTGTCAACAGCC 1444
QY 1394 GCTTCGAGCCCTGGCGCTGATCCAGCTGCACGCCATGCATACGTACGTACGTGCCCGTGG 1453
Db 1445 GCTTCGAGCCCTGGCGCTTCATCCAGCTCCAGGGAATGGCTATGGAACGCCGTGGTGT 1504
QY 1454 TAGCCTCCACCGCGCGCTGGTGCACACCGTCAAGGAGGCGCTACCGGCTTCCACATGG 1513
Db 1505 GCGGCTCCACCGCGCGGCTGCTGCACACGATCGTGGAGGCAAGACCTGGTTCACATGG 1564
QY 1514 GCGGCTGA-----ACCCGACAGCTGGACGAGGCTGACGCCGACGCCCTGGCGGCCA 1567
Db 1565 GCGGCTGACGCTGCACTGCAACGCTGGTGGAGCGCGGACGCTGAAGAAGGTGGCGACCA 1624
QY 1568 CGGTGCGCGTGCCACGAGGTGTTTGGCGGCGCGCTACCCCGAGATGGTGGCCAACT 1627
Db 1625 CCTGAGCGGCGCTCAAGGTGCTCGGCACGCGCGCGTACCAGGAGATGGTCAAGAACT 1684
QY 1628 GCATCAGCCAGGACCTGCTCTGCTCCAGCGCGCCAGAGTGGGAGGCGCTGCTGGAGG 1687
Db 1685 GCATGATCCAGGATCTCTCTGGAAGGACCTGCCAAGAACTGGGAGGAGGTGCTTCTGG 1744
QY 1688 AGTGTGTACGGCAAGGCGGCTGGCCACCGCCCAAGAGGAG 1731
Db 1745 AACTGGGGGTGGAGGGAGCGAGCGCGGGGATGCTCGCGGAGGAG 1788

Search completed: June 3, 2003, 13:33:39
Job time : 3720.76 secs

This Page Blank (uspto)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 05:08:24 ; Search time 308.686 Seconds
(without alignments)
15495.523 Million cell updates/sec

Title: US-09-980-771a-2

Perfect score: 2124

Sequence: 1 atggctgttgcctctaccag.....ccgcaccgcgaagtccgcc 2124

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	704	33.1	1696	22 AAC86955	CDNA encoding a gr
2	478	22.5	2267	18 AAX63355	Granule bound star
3	473.2	22.3	1915	19 AAX60319	DNA sequence of th
4	447.6	21.1	2542	19 AAV29753	Oryza sativa waxy
5	316.8	14.9	2161	24 AAK53210	Potato granule-bou
6	298	14.0	2115	21 AAK32824	Arabidopsis thalia
7	258	12.1	2183	24 AAK53215	Canna edulis granu
8	221	10.4	2807	21 AAC86435	Wheat starch synth
9	221	10.4	2842	21 AAC86411	Wheat starch synth

10	219.2	10.3	2097	19 AAV29755	Zea mays soluble s
11	218.2	10.3	2826	19 AAV01528	Wheat granule-boun
12	217.2	10.2	2423	19 AAV70958	DNA encoding maize
13	214	10.1	2107	21 AAC86412	Wheat starch synth
14	214	10.1	2107	21 AAC86434	Wheat starch synth
15	213.4	10.0	2380	19 AAV66834	Zea mays soluble s
16	212.4	10.0	2478	18 AAV67287	Soluble starch syn
17	212.4	10.0	9024	21 AAC86431	Wheat SSII gene SE
18	210.4	9.9	2920	24 AAK88112	CDNA encoding modi
19	209.2	9.8	2939	21 AAC86410	Wheat starch synth
20	209.2	9.8	2939	21 AAC86433	Wheat starch synth
21	208.8	9.8	2946	24 AAK88115	CDNA encoding barl
22	208.8	9.8	2951	24 AAK88114	CDNA encoding barl
23	208.2	9.8	2950	24 AAK88113	CDNA encoding barl
24	207	9.7	2248	21 AAZ50651	Corn starch synth
25	206.6	9.7	1798	21 AAZ50647	Corn starch synth
26	206.6	9.7	2019	21 AAZ50646	Corn starch synth
27	203.8	9.6	1724	24 ABA01821	Rice starch synth
28	161	7.6	2007	19 AAV70959	DNA encoding maize
29	161	7.6	2007	19 AAV29754	Zea mays soluble s
30	161	7.6	2085	18 AAT67286	Soluble starch syn
31	161	7.6	2085	19 AAV66833	Zea mays soluble s
32	151	7.1	1479	24 ABO90258	M. capsulatus gene
33	151	7.1	1479	24 ABO90406	M. capsulatus gene
34	125.6	5.9	333	21 AAV74569	Human ORFX ORF124
35	125.6	5.9	333	24 ABA18316	Human ORFX polynuc
36	118.2	5.6	2239	19 AAV01527	Wheat soluble star
37	118.2	5.6	2805	21 AAZ24487	Wheat soluble star
38	116	5.5	2662	20 AAX34651	CDNA sequence of w
39	110.8	5.2	65140	22 AAD17184	Streptomyces nous
40	110.8	5.2	125401	22 AAD17186	Streptomyces nous
41	110.6	5.2	2533	15 AQA45183	Soluble rice starc
42	108	5.1	1758	17 AAT32325	Soluble starch syn
43	106.8	5.0	2277	19 AAV13836	Human sapiens mamma
44	106.8	5.0	2277	19 AAV05372	Human telomerase p
45	106.4	5.0	1984	21 AAC46737	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAC86955

ID AAC86955 standard; cdna; 1696 BP.

AC AAC86955;

XX

XX

DT 02-APR-2001 (first entry)

XX

XX

DE CDNA encoding a granule bound starch synthetase II (GBSSII).

XX

XX

KW Granule bound starch synthetase II; GBSSII; starch grain;

KW adenosine diphosphate glucose-alpha,4-glucan alpha4-glucosyltransferase;

KW starch synthetase; ds.

XX

OS Chlamydomonas reinhardtii.

XX

XX

FT Key

FT CDS

FT

FT

FT

FT

XX

XX

PN FR2793806-AL.

XX

XX

PD 24-NOV-2000.

XX

XX

PF 21-MAY-1999; 99FR-0006494.

XX

XX

PR 21-MAY-1999; 99FR-0006494.

XX

XX

PA (CNRS) CNRS CENT NAT RECH SCI.

XX

XX

PI Dhulst C, Ball S;

XX WPI; 2001-052291/07.
DR P-PSDB; AAB311175.
XX
PT New recombinant nucleic acid encoding fusion of starch synthase and
PT second component, useful in pharmaceutical and food compositions, is
PT targeted to starch granules -
XX
XX
PS Claim 3; Fig 1; 52pp; French.
XX
XX The present sequence encodes a granule bound starch synthetase II
CC (GBSSI). The specification describes a recombinant nucleic acid,
CC comprising a polynucleotide encoding an adenosine diphosphate
CC glucose-alpha1,4-glucan alphas4-glucosyltransferase or starch
CC synthetase, placed upstream of a sequence that encodes a polypeptide
CC of interest. The adenosine diphosphate glucose-alpha1,4-glucan
CC alphas4-glucosyltransferase protein can migrate to sites of
CC biosynthesis of starch grains in plant cells, becoming associated
CC with these grains. The recombinant nucleic acid sequence is used to
CC target polypeptides of interest to starch grains.
XX
SQ Sequence 1696 BP; 347 A; 509 C; 571 G; 269 T; 0 other;

Query Match 33.1%; Score 704; DB 22; Length 1696;
Best Local Similarity 100.0%; Pred. No. 8e-94;
Matches 704; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1421 TGCACGCCATGCACTACGGTACCGTCCCGTGTAGCTCCACCGCGGCTGTGTGACA 1480
DB 13 TGCACGCCATGCACTACGGTACCGTCCCGTGTAGCTCCACCGCGGCTGTGTGACA 72

QY 1481 CCCTCAAGAGGGCGTCACGGCTTCCACATGGGCGCCCTGAACCCCGACAAGCTGAGC 1540
DB 73 CCCTCAAGAGGGCGTCACGGCTTCCACATGGGCGCCCTGAACCCCGACAAGCTGAGC 132

QY 1541 AGCTGACGCCGACGCGCTGGCGCCGACCGTGGCGCGTCCACGAGGTGTTTCGGGCG 1600
DB 133 AGCTGACGCCGACGCGCTGGCGCCGACCGTGGCGCGTCCACGAGGTGTTTCGGGCG 192

QY 1601 GCGCTACCCCGAGATGGTGGCCAACTGTCATCAGCCAGGACCTGCTCTGTCTCAAGCCCG 1660
DB 193 GCGCTACCCCGAGATGGTGGCCAACTGTCATCAGCCAGGACCTGCTCTGTCTCAAGCCCG 252

QY 1661 CCAGAGTGGAGGCGCTCTCGAGAGGTGTGTACGCAAGGGCGGCTGGCCACCG 1720
DB 253 CCAGAGTGGAGGCGCTCTCGAGAGGTGTGTACGCAAGGGCGGCTGGCCACCG 312

QY 1721 CCAGAGGAGGAGATCAAGTGCCTGTCGAGAGATCCCGGCGACCTGCGCCGCG 1780
DB 313 CCAGAGGAGGAGATCAAGTGCCTGTCGAGAGATCCCGGCGACCTGCGCCGCG 372

QY 1781 TGTCTACGCCCCCAACACCCCTGAAGCCCGTGTCCGCGCTCCGTGGAGGCAAGCGCGCG 1840
DB 373 TGTCTACGCCCCCAACACCCCTGAAGCCCGTGTCCGCGCTCCGTGGAGGCAAGCGCGCG 432

QY 1841 CCAGCGCCAGGTGCGGACACCGCCCGCCATGGCGCGTGGCGCGACACACCCCT 1900
DB 433 CCAGCGCCAGGTGCGGACACCGCCCGCCATGGCGCGTGGCGCGACACACCCCT 492

QY 1901 CGGGCCCTCGCCCGCGCGCCGACCCCAAGGTGACCACTACAAGCCCGCCCTGCGCG 1960
DB 493 CGGGCCCTCGCCCGCGCGCCGACCCCAAGGTGACCACTACAAGCCCGCCCTGCGCG 552

QY 1961 CCACCGCCAAAGCCGCTGCGCTCAAGTGGCGGTGAGGCGCTCCACCACTCGA 2020
DB 553 CCACCGCCAAAGCCGCTGCGCTCAAGTGGCGGTGAGGCGCTCCACCACTCGA 612

QY 2021 CCTCGGAGACGGCGCTGCGCTCAAGCGCAACGCAACGCAACGCTCGGCGCTCCAGACCT 2080
DB 613 CCTCGGAGACGGCGCTGCGCTCAAGCGCAACGCAACGCAACGCTCGGCGCTCCAGACCT 672

QY 2081 CGGTGCAAGCCCGCTGCTCGCGCCGACCCCGCAAGTCCGCG 2124
|||||

DB 673 CGCGTCCAAAGCCCGCTGGTCTCCGCGCCGCCACCCGCAAGTCGCGC 716

RESULT 2
AA63355
ID AAX63355 standard; cDNA; 2267 BP.
XX
XX AAX63355;
AC
DT 16-JUL-1999 (first entry)
XX
XX Granule bound starch synthase encoding cDNA.
DE
XX Maize; corn; Zea mays; delta-9 desaturase; GBSS; target; substrate;
KW granule bound starch synthase; hammerhead ribozyme; hairpin ribozyme;
KW modulation; gene expression; transgenic plant; cleavage; canola plant;
KW caffeine synthesis; coffee plant; nicotine production; tobacco;
KW fruit ripening; flower pigmentation; lignin production; ss.
XX
XX Zea mays.
XX
XX WO9710328-A2.
PN
XX 20-MAR-1997.
PD
XX 12-JUL-1996; 96WO-US11689.
PF
XX 13-JUL-1995; 95US-0001135.
PR
XX (DOWC) DOWELANCO.
PA
XX (RIBO-) RIBOZYME PHARM INC.
PI
XX Edington BE, Folkerts O, Guo L, McSwiggen JA, Merlo DJ;
PI Merlo PAO, Skokot TA, Young SA, Zwick MG;
XX
XX WPI; 1997-202224/18.
DR
XX Ribozyme which modulates plant gene expression - preferably
PT modulates expression of DELTA-9 desaturase or granule bound starch
PT synthase in maize or canola
XX
XX Example 9; Page 31-33; 155pp; English.
XX
XX The present invention describes an enzymatic nucleic acid molecule (I)
CC with RNA cleaving activity, which modulates the expression of a plant
CC gene. Also described is a gene comprising a cDNA sequence encoding maize
CC Delta-9 desaturase. (I) can be used to modulate expression of a gene,
CC preferably Delta-9 desaturase or a granule bound starch synthase (GBSS)
CC gene, in a plant (preferably a maize or canola plant). (I) can be used
CC to modulate caffeine synthesis in a coffee plant, nicotine production in
CC a tobacco plant, fruit ripening processes in an apple, tomato, pear,
CC plum or peach plant, flower pigmentation in a rose, petunia,
CC chrysanthemum or marigold plant or lignin production in a tobacco,
CC aspen, poplar or pine plant.
XX
SQ Sequence 2267 BP; 465 A; 699 C; 733 G; 370 T; 0 other;

Query Match 22.5%; Score 478; DB 18; Length 2267;
Best Local Similarity 59.8%; Pred. No. 4.4e-61;
Matches 986; Conservative 0; Mismatches 565; Indels 99; Gaps 7;

QY 161 GTGCCACCTTGGCGCTGGACATCGTGTGCTGCTGAGTCCGCGCTGGTCCCAAGA 220
DB 403 GCGCCACGCGCGGATGAAGCTGCTCTGCTGCGCGCGGAGATGGCGCGTGGAGCAAGA 462

QY 221 CGGGCGCGCTGGCGGATGTGACTGTGGCTGCTATTTAGCTGTGTAAGCGCGGCCACC 280
DB 463 CGGGCGCGCTGGCGGATGTGCTCTCGCGCGCTGCGCGCGCATGGCGCAATGGGCACC 522

QY 281 GCGTCATGACCAATTTGCGCTCGCTACGACCACTAGCTGACGCTGGGACACCTCGGTGG 340
DB 523 GTGTGATGGTGTCTCTCCCGCTACGACCACTACAGGACGCTGGGACACCTCGGTGG 582


```
Db 1544 CGACTGGTGGACTTGTGACACTGTGAAAGAGGCTACTACTGGATTCATATGGAGCCT 1603
QY 1520 TGA-----ACCCGCAAGCTGGACGAGGCTGACGGCGCCCTGGCCGCCACCCGTC 1573
Db 1604 TCAATGTTGAATGCGATGTTTGTGACCCAGCTGATGTGCTTAAAGATAGTAACAACAGTTG 1663
QY 1574 GCCGTGCCAGGAGGTGTTTCGGGGCCGCTACCCCGAGATGGTGCCTCAACTGCATCA 1633
Db 1664 CTAGAGCTCTTGGAGTCTATGGACCCCTCGCATTTGCTGAGATGATAAAAAATTCATGT 1723
QY 1634 GCCAGGACCTGTCTGTGTCCTCAAGCCGCCAGAAAGTGGAGGGCTGCTGGAGG 1687
Db 1724 CAGAGGAACCTCTCTGGAAGAACCTGCCAAGAAATGGAGACATTCCTATTGG 1777

RESULT 6
AAC32824
ID AAC32824 standard; DNA: 2115 BP.
XX
AC AAC32824;
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 779.
XX
DE Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
```


Db 1384 AGCAGAAGGGTCTTCATATCTAGTGGAGCTATTTCCAAAGTCTATGG---GGCTCAATG 1440
 QY 1235 TGCAGATGCCATCTCGTGGTACCGGCAAGCGCGCTACGAGAAGCTGGTGAACGCCATCG 1294
 Db 1441 TTCAATGGTTATCTCTGGGACTGGAAAGAAAGAAAGATGGAGCTCAGATTCTTGAACATG 1500
 QY 1295 GCACCAAGTACAAGGGCGCGCCAAAGGCGTGGTCAAGTTCTCGGCGCCCTTGGCGGACA 1354
 Db 1501 AAGAGAAGTCCAGGAGGCGGTGGAGTGGGGAATCAACGTGCGCATTTGGCTCAT 1560
 QY 1355 TGCTACCGCGCGCGCGCTCATCTGCTGGTCCCTCGCGCTTCGAGCCCTGCGGCCGGA 1414
 Db 1561 TGATCACTGTGGAGCTGACATTCATCTGCTCCCAAGCAGGTTTGAGCGGTGGTCTCA 1620
 QY 1415 TCCAGCTGCAGCGCATGCACTACCGTACCGTGGCGGTGGTACGCTCCACCGCGCGCTGG 1474
 Db 1621 TTCAGCTGCAGCGCATGAGATATGAAACCGCTCCCTATGTGGCATCTACTGGTGACTTG 1680
 QY 1475 TCGACACCGTCAAGGAGGCGCTACCGGCTTCCACATGGG-----CGCCCTGACCCCG 1528
 Db 1681 TGGACACTGTGAAGATGGCTACACAGTTTCCACATGGAGATTTCAACGTCAAGTGTG 1740
 QY 1529 ACAAGCTGGAGAGGCTGACGCCGACGCCCTTGGCGCCACCGTGGCGGTGCCAGCGAGG 1588
 Db 1741 AAGTTGGATCCAGATGATGTATAGCAACAGCAAGAGGCTGTGACAAAGCGGTTCGAG 1800
 QY 1589 TGTGTTGGGCGCGCTACCCGAGATGGTGGCCCAACTGATCAGCCAGGACCTGTCT 1648
 Db 1801 TATATGGAACATCCGAATGCAAGAAATGGTCAAGAACTGCATGGACCAAGACTTCTCT 1860
 QY 1649 GGTCCAAAGCGCGCCAGAAAGTGGAG 1674
 Db 1861 GGAAGGACCTGCGAGGTGTGGGAG 1886

RESULT 7

ABK53215

ID ABK53215 standard; cDNA; 2183 BP.

AC ABK53215;

XX

DT 14-AUG-2002 (first entry)

DE Canna edulis granule-bound starch synthase I (GBSSI) cDNA sequence.

XX Starch; amylose; cereal crop; commodity starch; thickened foodstuff;

KW bread, baking; pasta; potato; granule-bound starch synthase I; GBSSI; ss.

XX Canna edulis.

OS

XX W0200218606-A1.

PN

XX

PD 07-MAR-2002.

XX

PF 28-AUG-2000; 2000WO-US23494.

XX

PR 28-AUG-2000; 2000WO-US23494.

XX

PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX

PI Broglie K, Lightner J;

XX

XX WPI; 2002-404556/43.

DR

XX

XX

PT Novel starch isolated from cereal crop with a modification of amylose

XX content and/or amylose weight distribution relative to commodity starch

PT from same species of the cereal crop, for preparing thickened

XX foodstuffs

PS

XX Example 1; Page 37-38; 44pp; English.

XX

CC The present invention relates to a new starch comprising amylose,

XX where the starch is isolated from a cereal crop, and comprises a 50%

CC increase in average molecular weight relative to commodity starch from
 CC the same species of the cereal crop from which starch was isolated
 CC The starch of the invention is useful for preparing thickened foodstuffs
 CC by combining the starch, a foodstuff and water and cooking the resulting
 CC composition as necessary to produce a thickened foodstuff. The invention
 CC is also useful for preparing foodstuffs such as breads, baked goods,
 CC pastas, etc. The present nucleic acid sequence represents the Canna
 CC edulis granule-bound starch synthase I (GBSSI) cDNA sequence in clone
 CC ectlc.pk007.015 that was used in the methods of the invention for
 CC generation of a full-length sense Canna edulis GBSSI construct.

XX Sequence 2183 BP; 636 A; 428 C; 516 G; 603 T; 0 other;

Query Match 12.1%; Score 258; DB 24; Length 2183;

Best Local Similarity 51.6%; Pred. No. 3.4e-29;

Matches 794; Conservative 0; Mismatches 650; Indels 96; Gaps 5;

QY 176 TGGACATCGTGAATGTTGCTGCTGAGTGGCCCTTGGTCCCAAGAGCGGCGCTGGGCG 235

Db 357 TGAACCTTGGGGTTTGTGGTTGTAGTAGCTCCATGGAGCAAAAAGTGGGGCCTTGGCG 416

QY 236 ATGTGACTGGTGGCGCTGCTATTTAGCTGGTCAACGCGGCGCCACCGCTCATGACCATG 295

Db 417 ATGTCCTTATAGGATGGCCACTGCTATGGCTGCAATTTGGGCACAGGTCATGACCGTGG 476

QY 296 CCCCTCGCTACGACAGTACGCTGACGCTGGGACACCTCGGTGCTCGTGGACAT----- 350

Db 477 CGCCACGATATGACCAATATTAAGATATCTGGATACAAGTGTCCCACTTGAGTTAAAG 536

QY 351 -----CATGGCGGAGAGGTCCGCTACTTCCACTCCATCAAGAGGCGGTGACCGCG 403

Db 537 TTGGGATAGATTGAAACTGTCCGCTTCTTCCACTGTACAAAAGGGGAGTTGATCGGG 596

QY 404 TGTGATTGACCAACCCCTGTTCTCTGGCAAGCTCTGGGCAAGACCGGCTCCAGCTGT 463

Db 597 TTTTGTGATCACCCTATGTTCTCGAAGAGTTTGGGGAAACAGAGGAAAATAT 656

QY 464 AGGCCCCCGCTCGCGGCGCTGACTACTGAGCAACCAAGCGCTTCGCCCTGTCTGCA 523

Db 657 ATGGTCTGTTACAGGAACAGATTATGACAGACAATCACTAAGATTGACGCTTTTGTGCC 716

QY 524 AGGCGCTATGAGCTGCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576

Db 717 TGGCAGCTCTGGAAGCTCCAAAGACTTCTAAATCTCAACACAGCAAAATACTATTCTGAC 776

QY 577 -----GTCTTCTGCGCAACGACTGCTCCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 619

Db 777 CATATGAGATGATGTGTGTTTATTCGAACGATGGCATCTGCTCTACTGCCCTGCT 836

QY 620 TGCTGAAGGAGGATACAGCCCAAGGCGCCAGTTTCACCAAGCCCAAGTTCGGTGTGGCTA 679

Db 837 ACTTGAAGAACTATGTACCAATCATCATGTTTATGATGATGATGATGATGATGATGATGAT 896

QY 680 TCCACACATCGCCTTCCAGGCGCGCATGTTGGGAGGAGGCTTTCAAGGACACAGAGTGC 739

Db 897 TTCATAATATGCTTACCAGGCGCGCATTTGCTTTTCGAGCTTTTGAACCTCTTAACTCC 956

QY 740 CCCAGGCGCGCTTTGACAAGCTGGCCCTTCTGCGAGGCTATGCGCAAGTTCATGAGG 799

Db 957 CCATAAATTAAATCTTCATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1006

QY 800 CCACCCCATGGAGGAGGAG 859

Db 1007 -----GTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1025

QY 860 ACTGGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 919

Db 1026 ATTGGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1085

QY 920 CGACCGAGATGCTGCGCGATGCG 979

Db 1086 CCCAAGAGCTTGCTCAGGGGTAGAGAAGGAGTGTGAGGTTGGGCAATATCTCGCATGA 1145

QY 980 AGGCGATTGAGGCGATGTGAACGGCATGGACATTTGAGGAGTGGAAACCCCAAGACGACGA 1039
Db 1146 ARACCATCTGGAATAGTAAATGGATGGACACACCGAGTGGATCCATTAACAGACA 1205
QY 1040 AGTTCCTGCTGCGCCCTAGCAGCAGACAGCGCTCTACGCGCGCAAGCGCGCCCAAG 1099
Db 1206 AATATATTCTACAAACTACGATGCAACAACCTGATTGGATGCAAAACCTCTCTGTAAGG 1265
QY 1100 AGCCCTGAGCGGAGTGGGCTGCTGCTGGAACCCCGCCCTGCTGCGCCCTCA 1159
Db 1266 AAGCTTGGAAGCTGAGTGGGCTGCTGCTGTAACAAAACAAGCTGTTTGGGCTTGG 1325
QY 1160 TCGCGCCCTGAGGAGCAGAAAGGCTGTGACATCATCTCTGGCGCCCTGCCCAAGATCC 1219
Db 1326 TTGGAAGACTAGATGAGCAGAAAGGCTCAGACATTTAGTGCAGCAATTCAGAACTTC 1385
QY 1220 TGGCCACCCCAAGGTGAGATGCCATCTCTGGGTACCGGCAAGCGCCCTACGAGAAGC 1279
Db 1386 T--TTGTGAGAATGTTCAAGTGATAGTACTTGGCACTGGCAAGAGAAGTTGGAGAGTG 1442
QY 1280 TGGTGAACGCCATCGGCACCAAGTACAAAGGCGCGCAAGGCGTGGTCAAGTTCTCGG 1339
Db 1443 AACTTACATTACTTGAGGAATGTTTCCAGACAAATTCAGACACATCTCAAAATCAAG 1502
QY 1340 CGCCCTGCGGCACATGCTCACCGCGCGCGGACCTTCATGCTGGTGGCCCTCGCGCTCG 1399
Db 1503 TTCTTTAGCTCATGCAATCATGCGAGGAGCTGATATCTTGTATTCCAAAGCAGATCG 1562
QY 1400 AGCCCTGCGGCTGATCCAGCTGCAGCCATGACATACGCTACGCTACGCTGCGTGGTGCCT 1459
Db 1563 AACCCTGTGCGCTCATTTCAGCTTCAGGCCATGGGATATGGAACCTCTCCCTATGTTAGCA 1622
QY 1460 CCACGCGCGCTGCTGCGACACCTCAAGGAGGCGCTCACCGGCTTCCACATGGCGCC- 1518
Db 1623 CCACTGTGACCTGTGACACCTCAAGAGGCTTCAAGAGGCTTCACTGGCTTCCATATGGCCCT 1682
QY 1519 -----CTGAACCCCGCAAGCTGGAGGCTGAACGCGGCTGAGCGCGCTGCGCGCCACCGTGC 1573
Db 1683 TCAGTGTGGAGTGTGATGCGGTAGACAAAGCTGATGTACAAAGATTGTGGAACACGA 1742
QY 1574 GCGTGCACGAGGCTTTTGGGCGCGGCTACCCCGGAGATGTTGCGCAACTGCATCA 1633
Db 1743 AAGGGGCGCTCAAGTCTATGGAACACCTGCTTTTGGAGATGATCAAGAACTGCATGA 1802
QY 1634 GCACGACCTGCTGCTGCTCAAGCGCGCCCGCAGAAAGTGGGA 1673
Db 1803 ACCAAGATCTCATGGAAGGACCTGCAAGAAAGTGGGA 1842

RESULT 8
AAC86435
ID AAC86435 standard; cdna; 2807 BP.
XX
XX AAC86435;
AC
XX
XX 01-MAR-2001 (first entry)
DE
DE Wheat starch synthase clone wSSIID cdna.
XX
XX Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;
KW food product; adhesive; ss.
XX
XX Triticum aestivum.
XX
XX W0200066745-A1.
XX
XX 09-NOV-2000.
XX
XX 28-APR-2000; 2000W0-AU00385.
XX
XX 29-APR-1999; 99AU-0000052.
XX
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

PA (GOOD-) GOODMAN FIELDER LTD.
PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
XX
PI Morell M, Li Z, Rahman S, Appels R;
XX
XX WPI; 2000-647602/62.
XX
XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and
PT WST-II, useful in modifying plant starch content and/or composition -
XX
XX Disclosure; Fig 2; 211pp; English.
XX
XX The present invention relates to novel protein and coding sequences from
CC wheat. The proteins are wheat starch synthases, designated SSII and
CC SSIII. These can be used in the modification of plant starch content or
CC composition, and to screen plants to identify mutations which affect
CC starch content and composition. The starch can then be used in food
CC products, such as flour, and in films, coatings, adhesives, building
CC materials and packaging materials.
XX
SQ Sequence 2807 BP; 549 A; 844 C; 907 G; 507 T; 0 other;

Query Match 10.4%; Score 221; DB 21; Length 2807;
Best Local Similarity 56.2%; Pred. No. 7.6e-24;
Matches 485; Conservative 0; Mismatches 360; Indels 18; Gaps 3;

QY 851 AGAAGATCAACTGCTGAAGGGTGGCATTATCGCGCGCGCAAGCTGCTGCTGTCG 910
Db 1620 AGCAGCCCACTACTTTCGCGCGCGCTGAGATGGCGGACAGTTGCTGCTGAGGCC 1679
QY 911 CCAACTACGGGACGAGATCGCTCCGATGCCCGCGGGGTGTGGAGTGGACACCGTCA 970
Db 1680 CCGGCTACCTGTGGGAGCTCAAGCGGTGGAGGCGGCTGGGGCTTTCACGACATCATAC 1739
QY 971 TCC---CGCCAAAGGCGATTGAGGCGCATTTGAAAGCGATGACATTTGAGGAGTGGAAAC 1027
Db 1740 GGCAGAGCATGGAAGACCGCGCGCATTCGCAAGCGATCGCAACATGAGTGGAAAC 1799
QY 1028 CCAAGACCGACAAGTTCTGTCGCGCTTACGACACGAGAACAGCGTCT- 1075
Db 1800 CCGAGGTGGAGCTCCACCTCAAGTCGGACGCTACACCACTTCTCCCTGGGAGCGTGG 1859
QY 1076 ACGCCGCAAGCGCGCCCAAGGAGGCGCTGAGGCGGAGTGGCTGCTGTCGAC 1135
Db 1860 ACTCCGCAAGCGCGCTGCAAGGAGGCGCTGACGCGGAGTGGGCTGCGAGTTCGCG 1919
QY 1136 CCACGCGCGCTGCTGCTGCTTTCATCGCGCGCTTGGAGGAGCAGAAGGTTGGACATCA 1195
Db 1920 CCGAGTGGCGCTGCTGCTGCTTTCATCGCGCGCTTGGAGGAGCAGAAGGTTGGATCA 1979
QY 1196 TCCTGGCGCGCTGCGCCCAAGATTCCTGGCCACCCCAAGTGCAGATGCCATCTGGTA 1255
Db 1980 TCGCGGAGCGCATGCCCTGGATCGTG---AGCCAGAGCTGCTGCTGCTGCTGCGCA 2036
QY 1256 CCGCAAGCGCGCTTACGAGAGCTGTTGAGCGCATCGCACCAAGTACAGGCGCGCG 1315
Db 2037 CCGCGCGCGCATCGAGAGCTGCTGCGGCTTTCAGCGGAGGAGCACCACGACAGG 2096
QY 1316 CCAAGGCGCTGCTCAAGTTCGCGCGCGCTTGGGCGCATGCTACCGCGCGCGCGACT 1375
Db 2097 TCGCGGTTGGTGGGTTCTCCGTGCGCTTGGCGCACCGGATCAGCGCGCGCGCGACG 2156
QY 1376 TCATGCTGCTGCGCTGCGCTTTCGAGCGCTTGGGCTGATCAGCTGACGCCATGCACT 1435
Db 2157 CGCTCTCATGCCCTCCCGGTTTCGAGCGCTGCGGCTTCAACCGAGTTTACGCGCTTGGCT 2216
QY 1436 ACGTACCGTGGCGCTGAGCTCCACGCGCGCTGCTGACACCGCTCAAGGAGGCGG 1495
Db 2217 ACGGACCGTCCCGCTGCTGACGCGCTGCGGCGGTTGAGGAGCACCCTGCGCGCTTCG 2276
QY 1496 TCACCGCTTCCATGGCGCGCTTGAACCGCGCAAGCTGACGAGGCTGACCGCGAG 1555
Db 2277 ACCCGCTTCAACCACTCCGCGCTCGGGTGGACGTTGACCGCGCGCGCGAGGACAGCTGA 2336

QY 1556 CCCTGGCCGACACCGTGGCGGTGCGAGAGGTGTTTGGGGCGCGCTACCCGAGA 1615
 Db 2337 TCGAGGCCTCGGCGACTGCTCCGCACTACCGGACTACAAGAGAGCTGGAGGGCC 2396
 QY 1616 TGGTGGCAACTGCATCAGCAGGACCTGTCTGTCTCCAAAGCCCGCCAGAGTGGGAGG 1675
 Db 2397 TCCAGGACGGCGGATGTCGAGGACTTCAGCTGGGAGCATGCCGCCAAGCTCTACGAGG 2456
 QY 1676 GCCTGCTGGAGAGGTGTGTAC 1698
 Db 2457 ACGTCTCTCTCAAGGCCAAGTAC 2479

RESULT 9

AAC86411

ID AAC86411 standard; cDNA; 2842 BP.

XX AAC86411;

AC AAC86411;

DT 01-MAR-2001 (first entry)

XX Wheat starch synthase II coding sequence SEQ ID NO: 3.

XX Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;
 KW food product; adhesive; ss.

XX OS Triticum aestivum.

XX PN WO200066745-A1.

XX PD 09-NOV-2000.

XX 28-APR-2000; 2000WO-AU00385.

XX 29-APR-1999; 99AU-0000052.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PA (GOOD-) GOODMAN FIELDER LTD.
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.

XX PI Morell M, Li Z, Rahman S, Appels R;

XX WPI; 2000-647602/62.

DR P-PSDB; AAB37567.

XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and
 PT WST-II, useful in modifying plant starch content and/or composition -
 XX Claim 1; Page 155-158; 21pp; English.PS The present invention relates to novel protein and coding sequences from
 CC wheat. The proteins are wheat starch synthases, designated SSII and
 CC SSIII. These can be used in the modification of plant starch content or
 CC composition, and to screen plants to identify mutations which affect
 CC starch content and composition. The starch can then be used in food
 CC products, such as flour, and in films, coatings, adhesives, building
 CC materials and packaging materials.

XX

SQ Sequence 2842 BP; 561 A; 851 C; 919 G; 511 T; 0 other;

Query Match

DE Best Local Similarity 10.4%; Score 221; DB 21; Length 2842;

XX Mismatches 485; Conservative 0; Mismatches 360; Indels 18; Gaps 3;

QY 851 AGAAGATCAACTGGCTCAAGGGTGGGATTATCGCCCGCAGAAAGCTGGTGACTGTGCGC 910

Db 1620 AGCAGCGCAACTACTTCGCGCGCGGCTGTAGATGCGGACACAGTTGTCTGTGTAGGCC 1679

QY 911 CCAACTACGGACCGAGATGCTGCGATGCCCGCGGCTGTGGAGCTGGACACCGTCA 970

Db 1680 CCGGGTACCTGTGGGAGCTCAAGACGCTGGAGGGCGCTTCACGACATCATAC 1739

QY 971 TCC---GGCCCAAGGGCCATGTAGGGCATTTGAACGGCATGTGACATTTGAGGAGTGAACC 1027
 Db 1740 GGCAGAACACCTGTGAAGACCCGCGCATGTTCAAGGCATCGACAACATGTGAGTGAACC 1799
 QY 1028 CCAAGACCCGACAAGTTCTGTCTGCGCCCTAGCACAGAACACGGTCT-----1075
 Db 1800 CCGAGGTGACGCTCCACCTCAAGTCGAGCGGTACACCAACTTCTCCCTGGGAGCGCTGG 1859
 QY 1076 AGCCGCGCAAGCCCGCCCAAGGAGGCGCTGCAGCGGAGCTGGGCGCTGCTGTGGAC 1135
 Db 1860 ACTCCGGCAAGCGGCAGTTCAGAGAGGCGCTGCAGCGGAGCTGGGCGCTGCGAGTCCGG 1919
 QY 1136 CCACCGCCCGCTGTTCCGCTTCATCGCGCGCTTGGAGAGCAGAGGTGTGGACATCA 1195
 Db 1920 CCGAGCTGCGCTGCTCGCTTCATCGCGCGCTTGGAGAGCGGTGGAGATCA 1979
 QY 1196 TCCTGCGCGCGCTGCCAAGATCCTGGCCACCCCAAGGTGCAGATCGCCATCTCTGGGTA 1255
 Db 1980 TCGGAGCGCATGCCCTGGATCGT---AGCCAGGAGCTGCAGCTGTCTATCTGTGGCA 2036
 QY 1256 CCGCAAGCGCGCTTACGAGAAGCTGTGTGAACCGCATCGGCACCAAGTACAAAGGCCGCG 1315
 Db 2037 CCGCGCGCACGACCTGGAGAGCATGCTCGGCACTTCGCGGAGGACCAACGACACAGG 2096
 QY 1316 CAAGGGCGTGTCAAGTTCCTGGCGCCCTGTGGCGCACAATGCTACCCGCGCGCGACT 1375
 Db 2097 TGCOCGGGTGGGTGGGTTCCTCCGTGCGCTGGCGCACCGGATCACGGCGGCGCGCAGC 2156
 QY 1376 TCATGCTGTGCTCGCGCTTCGAGCGCTTGGGCGCTTGGGCGCTTACAGCTGCACGCCATGCACT 1435
 Db 2157 CGTCTCTATGCCCTTCCCGGTTTCGAGCGTTCGGGTTGAACCAAGCTTACGCCATGGCCT 2216
 QY 1436 ACGTACCGTCCCGTGTAGCTTCCACCGCGCGCTTGGTGTGACACCGCTCAAGAGGGCGC 1495
 Db 2217 ACGCACCGTCCCGTGTGCTGTCACGCGCTGCGGCGGTGAGGACACCTGCGCGCGTTCG 2276
 QY 1496 TCACGGGTTCACATGGGCGCTTGAACCCCGCAGCAAGCTGGAGAGCTGACGCCGACG 1555
 Db 2277 ACCCTTCAACCACTCCCGCTCGGGTGGAGCTTCGACCGCGCGGAGCACAAGCTGA 2336
 QY 1556 CCCTGGCGCGCACCGTGGCGGTGCCAGGAGGTGTTTGGCGGCGCGCTACCCCGAGA 1615
 Db 2337 TCGAGGCGCTCGGCACTGCTCCGCACTACCGGACTACAAGGAGAGCTGGAGGGGCC 2396
 QY 1616 TGTGGCAACTGCATCAGCAGGACCTGTCTGTGTCAAGCCCGCCAGAAAGTGGGAGG 1675
 Db 2397 TCCAGGAGCGGCGATGTCGAGGACTTCAGCTGGGAGCATGCCGCCAAGCTCTACGAGG 2456
 QY 1676 GCCTGTGGAGGAGGTGTGTAC 1698
 Db 2457 ACGTCTCTCTCAAGGCCAAGTAC 2479

RESULT 10

AAV29755

ID AAV29755 standard; DNA; 2097 BP.

XX AAC29755;

XX 11-SEP-1998 (first entry)

DE Zea mays soluble starch synthase IIb gene.

KW SER; starch-encapsulating region; fusion vector;

KW soluble starch synthase IIb; glucosyl transferase; ss.

XX Zea mays.

XX Key

FH 1.2097

FT CDS

FT /tag= a

FT /product= soluble starch synthase IIb

XX Location/Qualifiers

PN W09814601-A1.
 XX 09-APR-1998.
 XX 30-SEP-1997; 97WO-US17555.
 XX 30-SEP-1996; 96US-0026855.
 XX (EXSE-) EXSEED GENETICS LLC.
 XX Guan H, Keeling P;
 XX WPI: 1998-240100/21.
 DR P-PSDB; AAW56487.
 XX Hybrid polypeptide comprising starch-encapsulating region and
 PT protein - useful for, e.g. producing protein(s) resistant to
 PT degradation by stomach acids
 XX Example 2; Page 36-38; 156pp; English.
 XX The sequence is that of the soluble starch synthase IIb gene.
 CC It can be used in the production of a hybrid polypeptide
 CC comprising a starch-encapsulating region (SER) fused
 CC to a payload protein. The hybrid polypeptide can be used to make
 CC modified starches comprising the payload protein, selected from,
 CC e.g. hormones, growth factors, antibodies, enzymes, dyes,
 CC immunoglobulins, etc. The modified starch can also be used
 CC to provide grain feeds enriched in amino acids. By encapsulating
 CC the payload protein in starch, it is more resistant to
 CC degradation by stomach acids.
 XX Sequence 2097 BP; 409 A; 575 C; 696 G; 417 T; 0 other;
 SQ

Query Match 10.3%; Score 219.2; DB 19; Length 2097;
 Best Local Similarity 50.8%; Pred. No. 1.4e-23;
 Matches 801; Conservative 0; Mismatches 673; Indels 108; Gaps 7;

QY 144 TGTACTGGTGCACCTGGTGCACCTGGCGCTGGACATCGTATGTTGCTGAGGT 203
 DB TGAACCTGGCCCTTGGCTGGCCCTAATGTATGAACGTGCTGCTGGCTTCTGAATG 647
 QY 204 CGCCCTTGGTCCAAAGCGGCGCTGGGGATGCTGACTGGTGGCTGCTGCTATGAGCT 263
 DB TGTCTCTTCTGCAAGACAGGTGGCTTGGAGATGCTGGTGGTCTTGGCTTAAGGCTCT 707
 QY 264 GGTCAAGCGGCGCACCGCTCATGACCATTTGCCCTCGCTACGACCAAGTACGCTGACGC 323
 DB GGCAGAGGAGGACACCGTGTATGTTGCTGATACCAAGATATGAGAGTATGCCGAGC 767
 QY 324 CTGGGACACCTCGGTGCTGGACATCATGGCGGAGAAAGTCCGCTACTTCCACTCCAT 383
 DB CCGGG--ATTAGTGTAAAGGAGACGTTACAAGCTAGCTGACACAGGATTCAGAAGTTACT 825
 QY 384 CAAGAAGGGGTGCACCGGTGGATGTGACCCCTGTTCTGTCGCGCAAGTCTGGGG 443
 DB TAITTTCACTCTTACATTTGATGAGTGTATTTCTATTTCTAGAAAGCCCTCCCTTCGGG 885
 QY 444 CAAGACCGGCTCCAAAGCTGTACGGCCCGCTCGGGCTGACTACCTGGACACCAACAA 503
 DB 886 CAC-----CGGCACAATATATTTATGGGGGAGAAAGATTTGATATTTGAA 932
 QY 504 GCCTGTCCTGCTGTCAGGCGCGCTATTTAGGCTGCGCGGTGCTGCGCTTCGGGCC 563
 DB 933 GCGCATGATTTGTTCTCAAGGCGCTGTTGAGTGTCCATGGTATGCTCCATGTGGCGG 992
 QY 564 -----CGGCGAGGACTGGTCTTCTGGCCAAACACTGGCACTCGGCCCTGGT 611
 DB 993 TACTGTCTATGATGGAACACTAGTTTTCATGCTAATGATTGGCATACCGCACTTCT 1052
 QY 612 GCGCTGCTGCTGAAGGACGATACCAAGCCCAAGGCGCAGTTCACCAAGGCCAAGTCCGT 671
 DB 1053 CCCTGTCTATCTAAGGCTTATCCGGGACAAATGTTGATGTCAGTATGCTCGCTCTGT 1112

QY 672 GCTGGCTATCCACAACATCGCTTCCAGGCGCCGATGTGGAGGAGGCTTTCAGAGACAC 731
 DB 1113 GCTTGTATACACAACATTCCTCATCAGGCTGCTGGCCCTGTAGACGACTTCGTAATTT 1172
 QY 732 GAAGCTGCCCGAGGCGCCCTTTGACAAGCTGGCCCTTTCGACGGGCTATCCCAAGGTTTA 791
 DB 1173 TGACTTGCCTGAACACTATACATCGACCACTTCAAACTGTATGACAACATTTGGTGGGATCA 1232
 QY 792 CACTGAGGCCACCCCATGGAGGAGGACGAGAACCCCGCTGACGGGAAAGACCTACAA 851
 DB 1233 C----- 1233
 QY 852 GAAGATCAACTGGCTGAAGGTTGGCAATTTATCGCGCCGACAAAGCTGTGTGTCGCC 911
 DB 1234 ----AGCAACGTTTTTCTGCGGGCTGAAGACGGCAGACCGGGTGTGACCGCTTAGCAA 1289
 QY 912 CAACTAGCGACCGAGATCGCTGCCGATGCCCGCGGGTGTGAGCTGGACACCGCTCAT 971
 DB 1290 TGGCTACATGTGGGAGCTGAAGACTTCGSAAGGCGGGTGGGCCCTCCACGACATATAA 1349
 QY 972 CC---GGCCCAAGGCGCATTTGAGGCGCATTTGAACGCGCATGGACATTTGAGGAGTGAACCC 1028
 DB 1350 CCAGAAGCACTGGAAAGCTGCAGGSCATCGTGAACGSCATCGACATGACGAGTGAACCC 1409
 QY 1029 CAAGACCGACAAGTTCCTGCTGTGCGCCCTAGCACCGACAACAGC-----GTCTA 1076
 DB 1410 CGCTGTGACGTGCACCTCCACTCCGACGACTACACAACTATACACTTCGAGACGCTGGA 1469
 QY 1077 GCGCGGCAAGGCCCGCCAGGAGGCGCTGCAGGCGGAGCTGGCCCTGCTGCTGTGGACCC 1136
 DB 1470 CACCGCAAGCGGAGTGAAGGCCCGCTGCAGCGGCGCTGGGCCCTGCAAGTCCCGGA 1529
 QY 1137 CACGCGCCCTGTTTCGCTTCATCGCGCGCTGGAGGAGCAGAGGCTGTGGACATCAT 1196
 DB 1530 CGAGCTGCCACTGATCGGTTTCATCGGCGGCTGGACCAACAGAGGCGCTGGACATCAT 1589
 QY 1197 CCTGGCGCCCTGCCAAGATCTTGGCCACCCCAAGTGCAGATCGCCATCTCTGGGTAC 1256
 DB 1590 GCGCGAGCGGATCCACTGGATC---GCGGGGCAAGACGCTGCACTGCTGTGCTGGGAC 1646
 QY 1257 CGCAAGGCGCCCTACGAGAAGCTGTTGAACGCCATCGGCACCAAGTACAAGGCGCGCGC 1316
 DB 1647 CGGCGGCGCCCACTGGAGGACATGCTCGCGGGTTCAGTCCGNGCACAGGACAAGGT 1706
 QY 1317 CAAGGCGGTGTGAAGTTCTCGGCGCCCTGGGCGCATGCTACCGCGCGCGCGGACTT 1376
 DB 1707 GCGCGCGTGGGTTCGTTCCGTTCCGCTGGCGCACCGCATACGGCGGCGCGGACAT 1766
 QY 1377 CATGCTGTGCGCTCGCGCTTCGAGCCCTCGGCGCTGATCCAGCTGCACGCTACCTA 1436
 DB 1767 CTTGCTGATGCGTGGCGCTTCGAGCGCTGGGGCTGAACCAAGCTCTACGCCATGGCGTA 1826
 QY 1437 CGGTACCTGCGCGTGTAGCTTCCACCGCGGCTGCTGTCGACACCGCTCAAGGAGGCGGT 1496
 DB 1827 CGGACCGTGGCGGTGTCAGCGCTGGGGGCTCCGCGACAGGCTGGCGCGCTTCA 1886
 QY 1497 CACCGGCTTCCACATGGGCGCCCTGAACCCCGACATGCTGACGAGGCTGAGCGCGAGC 1556
 DB 1887 CCGCTTCAACGACACCGGCTCGGTTGGACGTTTCACGCGCGGAGGCGGACCGATGAT 1946
 QY 1557 CTTGGCGCCCGCTGGCGCTGCAGCGAGGTGTTTGGGGCGCGCTACCCCGAGAT 1616
 DB 1947 CGAGCGCTCTCGCACTGCTTCCACCGCTACCGGAACTACAAGGAGAGCTGCGCGCGCTG 2006
 QY 1617 GGTGGCAACTGCTACGACGAGGACCTGCTCTGTGTGTCACAGCCCGCCAGAGTGGAGGG 1676
 DB 2007 CAGGCGCGCGCATGGCGGAGGACCTCAGCTGGGACCAACGCGCGCTGTATAGGA 2066
 QY 1677 CTTGCTGAGGAGGTGGTGTAC 1698
 DB 2067 CTTGCTGCAAGGCGAAGTAC 2088


```
Db 1234 ----AGCAACGTTTTTCTGGGGGTGAAGACGGACGACCGGGTGGTACCGGTAGCAA 1289
QY 912 CAACCTAGCGACCGAGATCGTCCGATGCGCGGGGGTGTGAGCTGGACACCGTCAT 971
Db 1290 TGGCTACATGTGGGAGTGAAGACTTCGGAAGGCGGGTGGGCTCCACGACATCATAAA 1349
QY 972 CC---GCGCAAGGCGATTAAGGSCATGTGAACGGCATGCACATTAAGAGTGAACCC 1028
Db 1350 CCAGAACGACTGNAAGTGCAGGCGATCGTGAAGCGCATGCACATGAAGAGTGAACCC 1409
QY 1029 CAGACCGACAAGTTCCTGTCTGGCCCTACGACCAAGAACAGC-----GTCTA 1076
Db 1410 CGCTGTGAGCTGCACCTCCACTCCGACGACTACACCAACTACACGTTGAGAGCGTGA 1459
QY 1077 CGCGGCAAGGCGCCGCCAAGAGAGGCCCTGACGCGCGAGCTGGGCTGCTGTGGAACCC 1136
Db 1470 CACGGCAAGGCGAGTGCAGAGCGCGCCCTGCAGCGCAGCTGGGCTGCAGGTCCGCGA 1529
QY 1137 CACGCGCCCGCTGTTCGCTTCATCGSCGCGCTGGAGGAGCAGAAAGGTGTGACATCAT 1196
Db 1530 CGAGTGCCTCATGATCGGTTTCAITGGCGGCTGGACCCAGAAAGGCGTGGACATCAT 1589
QY 1197 CTGGCGCGCTGCCAAGATPCTGGCCACCCCAAGGTGCAGATCGCCATCCTGGGTAC 1256
Db 1590 CGCGGAGCGATCCACTGGATC---GCGGGCAGGAGCTGCAGCTSGTATGCTGGCAC 1646
QY 1257 CGCMAAGCGCGCTACGAGAGTGTGAACGCATCGGCAACCAAGTACAAAGGCGCGCG 1316
Db 1647 CGGCGGCGCGCTGAGAGACATGTGCGCGCGTTCAGTCTGAGCAGCACAGCAAGGT 1706
QY 1317 CAAGGCGGTGTCAAGTTCCTGGCGCGCCCTGGCGGCACATGCTCACGCGCGCGCGACTT 1376
Db 1707 GCSCGCTGGTGGGTCTCTCGTGGCCCTGGCGCACCGCATCACGGGCGCGGACAT 1766
QY 1377 CATGCTGGTCCCTCGCGTTCGAGCCCTGCGGCCCTGATCCAGCTGCACGCCATGCACTA 1436
Db 1767 CTGCTGATGCGCTGCGGTTTCGAGCGCTGCGGGCTGAACCACTCTACGCCATGGCGTA 1826
QY 1437 CGGTACCGTGGCGTGTAGCTPCCACCGCGCGCTGGTTCGACACCGTCAAGAGGCGGT 1496
Db 1827 CGGACCGTGGCGTGTGACGCGCTGGGGGGCTCCGGGACACAGGTGGCGCGCTTCA 1886
QY 1497 CACCGCTTCCATATGGGCGCTGTAAACCCCGACAAGCTGGAGAGCTGACGCCGACGC 1556
Db 1887 CCGCTTCAACGACACCGCGCTGCGGTGGAGTTCGACCGCGCGGAGCGAACCGATGAT 1946
QY 1557 CTTGGCGCGCACCGTGGCGCTGCCAGCGAGTGTGTTGGGGGCGCGCTTACCCCGAGAT 1616
Db 1947 CGACGCGCTCTCGCACTGCTCACCACGTACCGGAACCTACAAAGGAGAGCTGGCGCGCTG 2006
QY 1617 GGTGGCCAACTGCATCAGCCAGGACCTGTCTGGTCCAGCGCGCGCCAGAGTGGGAGG 1676
Db 2007 CAGGGCGCGCGCTGGCGGAGACCTCAGCTGGGACCAAGCGCGCTGTGTATAGGA 2066
QY 1677 CTGTGTGAGGAGGTGGTGTAC 1698
Db 2067 CTGTGTGCTCAAGGCGAAGTAC 2088
```

RESULT 13

AAC86412

ID AAC86412 standard; cDNA; 2107 BP.

XX AAC86412;

XX AAC86412;

XX AAC86412;

DT 01-WAR-2001 (first entry)

XX Wheat starch synthase II coding sequence SEQ ID NO: 5.

DE Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;

KW food product; adhesive; ss.

XX Triticum aestivum.

OS Triticum aestivum.

```
XX WO200066745-A1.
PN 09-NOV-2000.
XX 28-APR-2000; 2000WO-AU00385.
XX 29-APR-1999; 99AU-0000052.
XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA (GOOD-) GOODMAN FIELDER LTD.
PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
XX Morell M, Li Z, Rahman S, Appels R;
XX WPT; 2000-647602/62.
XX P-PSDB; AAB37568.
XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and
XX WST-II, useful in modifying plant starch content and/or composition -
XX Claim 1; Page 161-163; 211pp; English.
XX The present invention relates to novel protein and coding sequences from
XX wheat. The proteins are wheat starch synthases, designated SSII and
XX SSIII. These can be used in the modification of plant starch content or
XX composition, and to screen plants to identify mutations which affect
XX starch content and composition. The starch can then be used in food
XX products, such as flour, and in films, coatings, adhesives, building
XX materials and packaging materials.
XX Sequence 2107 BP; 452 A; 566 C; 664 G; 425 T; 0 other;
XX Query Match 10.1%; Score 214; DB 21; Length 2107;
XX Best Local Similarity 55.8%; Pred. No. 8e-23;
XX Matches 478; Conservative 0; Mismatches 360; Indels 18; Gaps 3;
```

```
QY 858 CAACCTGCTGAAGGTGGCAATTATCGCGCGCAGCAAGCTGCTGCTGCGCCAACTA 917
Db 933 CAACCTACTTCCGCGCGCTGAAGATGCGGACAGGTTGCTGTGAGCCCGCGGTA 992
QY 918 CGCGACCGAGATCGTCCGATGCGCGCGCGGTGTGGAGCTGGACACCGTCATCC---G 974
Db 993 CTGTGGGAGCTGAAGACCGGTGGAGGCGCTTTCAGGACATCATACGGCAGAA 1052
QY 975 CGCCAAAGGGCATTTGAGGCGCATTTGAAACGCATGGACATTTGAGAGTGAACCCAGAC 1034
Db 1053 CGACTGGAAGACCGCGGCGATCTCAACGCGATCGACAGCATCGACAACTACATCGGAGGT 1112
QY 1035 CGACAAGTTCCTGCTGCGCCCTTACGACCAAGACAGCGTCT-----ACGCGCG 1082
Db 1113 GGACGCGCCACTCAAGTCGGAGCGCTTACCAACTTCTCCCTGAGGACGCTGACTCCGG 1172
QY 1083 CAAAGCGCGCGCAAGGAGGCGCTTGCAGGCGGAGCTGGGCTGCTGCTGAGACCCACCGC 1142
Db 1173 CAAAGCGCGAGTGCAGAGGCGCGCTTGCAGGCGGAGCTGGGCTGCTGCTGAGTCCGCG 1232
QY 1143 CCCCCTGTTCTGCTTTCATCGCGCGCTTGCAGGAGCAGAGGCTGTGAGCATCATCTGGC 1202
Db 1233 GCGCTGCTGCTTTCATCGCGCGCTTGCAGGCGCAGAGGCGCTGGAGATCATCGCGA 1292
QY 1203 CGCGCTGCGCAAGATCCTTGCACCCCAAGGTGCGAGATCGCCATCCTGGTACCGGCAA 1262
Db 1293 CGCCATGCGCTGGATCGTG---AGCCAGGACGCTGACGCTGATGCTGGGACCGGCG 1349
QY 1263 GCGCGCTACGAGAGCTGTTGAACCCATCGCACCAAGTACAGGCGCGCGCAAGGG 1322
Db 1350 CCAGACCTTGAGAGAGCTGCTGACGACTTCGAGCGGAGCAGCACAGCAAGGTGCGCG 1409
QY 1323 CTGTGCTCAAGTTCCTGCGCGCGCTTGGCGCACATGCTCACCGCGCGCGGCTTATGCT 1382
Db 1410 GTGGGTGGGTCTCTCGTGGCGCTTGGCGCACCGGATTCAGGCGGGGCGGACGCGCTCT 1469
```

QY 1383 GGTGCGCTCGCGCTTCGAGCCCTCGCGCTGATCCAGCTCCAGCCATGCACTACGGTAC 1442
 ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1470 CATGCCCTCCCGGTTGCTGGCTGGGGCTGACACAGCTTACGCCATGCGCTACGGCAC 1529
 ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1443 CGTGGCCGTTAGCTCCACCGCGCGCTGTCGACACCGCTCAAGAGAGGCGCTACCGG 1502
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1530 CGTCCCGCTGTCACGCGCTCGCGCGCTCAGGGACACCGTCCCGCTTCGACCCCTT 1589
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1503 CTTCCACATGGCGCCCTGACCCCGACAGCTGGAGGAGCTGACCGCGCCCTGGC 1562
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1590 CAACCACTCCGGCTCGGGTGGAGCTTCGACCGCGCGGAGGCGCAACAGTGTATCGAGG 1649
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1563 CGCCACCGTCCGCTGCGACGAGGTGTTTGGCGGCGCGCTACCCCGAGATGTGGC 1622
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1650 GCTCGGCACTGCTCCGCACTTACGAGACTTCAAGGAGAGCTGGAGGSCCCTCCAGGA 1709
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1623 CAACCTGATCAGCAGAGCTGCTGCTGCTCCAGCCCGCGCCAGAGTGGAGGCGCTGCT 1682
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1710 GCGCGGATGTCGAGGACTTACGCTGGGAGCACGCGCGCAAGCTCTACGAGGACGCTT 1769
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1683 GGAGGAGTGGTGTAC 1698
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1770 CGTCAAGGCCAAGTAC 1785
 ||||| || || || || || || || || || || || || || || || || || || ||

RESULT 14

AAC86434
 ID AAC86434 standard; cDNA; 2107 BP.

XX AAC86434;

XX 01-MAR-2001 (first entry)

XX Wheat starch synthase clone wSIIIB cDNA.

XX Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;
 KW food product; adhesive; ss.

XX Triticum aestivum.

XX WO20006745-A1.

XX 09-NOV-2000.

XX 28-APR-2000; 2000WO-AU00385.

XX 29-APR-1999; 99AU-0000052.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX (GOOD-) GOODMAN FIELDER LTD.

XX (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.

XX Morell M, Li Z, Rahman S, Appels R;

XX WPI; 2000-647602/62.

XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and
 WST-II, useful in modifying plant starch content and/or composition -

XX Disclosure; Fig 2; 21lpp; English.

XX The present invention relates to novel protein and coding sequences from
 CC wheat. The proteins are wheat starch synthases, designated SSII and
 CC SSIII. These can be used in the modification of plant starch content or
 CC composition, and to screen plants to identify mutations which affect
 CC starch content and composition. The starch can then be used in food
 CC products, such as flour, and in films, coatings, adhesives, building
 CC materials and packaging materials.

XX Sequence 2107 BP; 452 A; 566 C; 664 G; 425 T; 0 other;

XX Query Match 10.1%; Score 214; DB 21; Length 2107;

XX Best Local Similarity 55.8%; Pred. No. 8e-23;

Matches 478; Conservative 0; Mismatches 360; Indels 18; Gaps 3;
 QY 858 CAACTGCTCAAGGGTGGCAATTATCGCCGCCGACAAAGCTGGTGAATGTGCGCCAACTA 917
 ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 933 CAACTACTTCGCCCGCGGCTGAAGATGGCGACCAAGTTGTGTGTGAGCCCGGGTA 992
 ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 918 CGGCAACGAGATCGTCCCATGCGCGCGGCTGGAGTGGAGTGGACACCGTATCC---G 974
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 993 CTTGTGGAGCTGAAGACGCTGGAGGCGCTGGGGCTTCACGACATCATACGGCAGAA 1052
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 975 CGCAAGGGATTCAGGGCATTTGAACGCGCATGGACATTGAGGATGGAAACCCCAAGAC 1034
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1053 CGACTGGAAGACCGCGGCATCGTCAACGCGATCGACAAATGAGTGGAAACCCGAGGT 1112
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1035 CGACAAGTTCTGTCTCGCCCTACGACACAGACAGTCT-----ACGCCG 1082
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1113 GGAGCCCACTCAAGTCGGACGGCTACCAACTTCTCCCTGAGGACGCTGGACTCCG 1172
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1083 CAAGCGCCCGCAGAGGCGCTCGAGCGCGAGCTGGGCTGCTGTGGACCCACCGC 1142
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1173 CAAGCGGACGTGCAAGAGGCGCTGCGCGAGCTGGGCTGCGAGTTCGCGCGGACGT 1232
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1143 CCCTCTGTTCGCTTCATCGCGCGCTGGAGGACAGAAGGTTGGACATCATCTTGGC 1202
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1233 GCGCTGCTCGCTTCATCGCGCGCTGGACGGCGTGGAGGCGTGGAGATCATCGGGA 1292
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1203 CGCCCTGCCCAAGATCTTGGCCACCCCAAGTGCAGATGCCATTCCTTGGTACCGGAA 1262
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1293 CGCCATGCCCTGGATCGTG---AGCCAGGACGTGCAGTGGTGTGCTGGGACCGGCG 1349
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1263 GGCGGCTACAGAGAGCTGGTGAACGCCATCGGCACCAAGTACAAAGGCGCGCCCAAGG 1322
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1350 CCACGACCTGGAGAGCATGCTGCAGCACTTCGAGCGGAGCACCAAGAGGTGCGCGG 1409
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1323 CGTGTCAAGTTCTCGCGCGCTGGCGCACATGCTACCGCGCGCGCGGCTTCACTGCT 1382
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1410 GTGGTGGGTTCTCCGTCGCGCTGGCCACCGATACGCGGGGCGGAGCGCTCT 1469
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1383 GGTGCTCGCGCTTCGAGCGCTTGGGCTGATCCAGTGCAGCGCATGCACTACGGTAC 1442
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1470 CATGCCCTCCCGTTCTGTCGCGCTGCGGCTGAACCAAGCTTACGCGCATGGCTACG 1529
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1443 CGTGGCGTGTAGCTCCACCGCGGCTGTCGACACCGCTCAAGGAGGCGTCAACGG 1502
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1530 GTCCCGCTGCTGCACGCGCTCGCGCGCTCAGGACACCGCTGCGCGCTTCGACCCCTT 1589
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1503 CTTGCACATGGCGCGCTGAACCGCGCAAGCTGGACAGGCTGACGCGCGGCGCTGCG 1562
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1590 CAACCACTCCCGGCTCGGTTGAGCTTCGACCGCGCGGAGGCGCACAGCTGATCGAGG 1649
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1563 CGCCACCGTGGCGCTGCCAGCGAGGTGTTTGGCGGCGCGGCTTACCCGAGATGGTGGC 1622
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1650 GCTGGGCACTGCGCTCCGCACTACCGAGACTTCAAGGAGAGCTGGAGGCGCTCCAGGA 1709
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1623 CAACTGATCAGCAGGAGCTTCTGCTCAAGCCCGCCAGAGTGGAGGCGCTGCT 1682
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1710 GCGGCGCATGTCGAGGAGCTTACGCTGGGAGCACCGCGCGCAAGCTCTACGAGGAGCTCT 1769
 ||||| || || || || || || || || || || || || || || || || || || ||
 QY 1683 GGAGGAGTGGTGTAC 1698
 ||||| || || || || || || || || || || || || || || || || || || ||
 Db 1770 GGTCAAGGCCAAGTAC 1785
 ||||| || || || || || || || || || || || || || || || || || || ||

RESULT 15

AAV66834
 ID AAV66834 standard; DNA; 2380 BP.

XX AAV66834;

XX 05-JAN-1999 (first entry)

XX Zea mays soluble starch synthase gene SSS56.

XX

Db 2141 ARTAYCARTGGTTRGCGNAAATRYTNGCNACNMGNGNWSNTGYMGNMGNACTGGA 2200
QY 1752 CGAGAAAGATCCCCGGCGACCTGCCCCCGGTGCTCTACGCCCCCAACACCCCTGAAGCC 1808
Db 2201 CNYTNTTYMGNMGNYTNTTYSNYTNGCNCNYTNTATGMGNGCNSNCAYTTNNGNM 2257

Search completed: June 3, 2003, 10:07:07
Job time : 351.686 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 09:33:29 ; Search time 2072.3 Seconds
(without alignments)
16599.542 Million cell updates/sec

Title: US-09-980-771a-2

Perfect score: 2124

Sequence: 1 atgctgttgctctaccag.....cgcaccgcgaagtcgcgc 2124

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_nam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	508.6	23.9	786	12	BF864001 963048D01
2	503.6	23.7	511	10	AV394077 AV394077
3	476.6	22.4	502	14	B0814687 1030045B1
4	469.8	22.1	473	10	AV640741 AV640741
5	466.8	22.0	506	10	AV643281 AV643281
6	458	21.6	475	10	AW757933 874004E12

7	457.8	21.6	461	10	AV629488
8	456.8	21.5	2147	11	AY109531
9	456	21.5	666	13	BM003295
10	444.4	20.9	458	10	AV622787
11	443.8	20.9	560	10	AV628313
12	432.8	20.4	436	10	AV629117
13	422.8	19.9	552	10	BE024926
14	421.2	19.8	517	10	AV641583
15	415.2	19.5	532	10	AV641989
16	414.8	19.5	418	10	AV644765
17	406.8	19.2	513	10	AV631004
18	406.8	19.2	552	10	AV640546
19	405.8	19.1	524	10	AV641724
20	405.2	19.1	500	10	AV642869
21	402.8	19.0	510	10	AV395307
22	394.8	18.6	540	10	AV642722
23	391.4	18.4	393	10	AV631115
24	382.8	18.0	486	10	AV642834
25	382.8	18.0	504	10	AV629198
26	381.8	18.0	502	10	AV644517
27	379.8	17.9	495	10	AV642565
28	371.8	17.5	493	10	AV631838
29	370.8	17.5	492	10	AV629891
30	369.8	17.4	484	10	AV629266
31	369.8	17.4	489	10	AV629190
32	364.8	17.2	487	10	AV643763
33	364.6	17.2	385	10	AV393664
34	359.8	16.9	490	10	AV628140
35	358.8	16.9	476	10	AV644278
36	356.8	16.8	478	10	AW758015
37	356.8	16.8	502	10	AV631098
38	354.8	16.7	474	10	AV397080
39	353.8	16.7	468	10	AV644593
40	352.8	16.6	474	10	AV630110
41	352.6	16.6	367	13	BI727902
42	351.8	16.6	487	10	AV629387
43	351.2	16.5	799	12	BF863935
44	330.2	15.5	453	10	AV643760
45	326.8	15.4	448	10	AV644481

ALIGNMENTS

```

RESULT 1
BF864001
LOCUS
DEFINITION
  BF864001
  786 bp mRNA linear EST 19-JAN-2001
  C. reinhardtii CC-1690, Stress condition I, normalized
  , Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION
  BF864001
VERSION
  BF864001.1 GI:12254145
KEYWORDS
  EST.
SOURCE
  Chlamydomonas reinhardtii.
ORGANISM
  Chlamydomonas reinhardtii.
  Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
  Chlamydomonadales; Chlamydomonas.
REFERENCE
  1 (bases 1 to 786)
AUTHORS
  Grossman, A., Davies, J., Federspiel, N., Harris, E., Hauser, C.,
  Lefebvre, P., McDermott, J. P., Shrager, J., Silflow, C. and Stern, D.
  Analyses of the Chlamydomonas reinhardtii Genome: A Model,
  Unicellular System for Analyzing Gene Function and Regulation in
  Vascular Plants; project phase 3
JOURNAL
  Unpublished (2000)
COMMENT
  Contact: Charles Hauser
  DCMB Box 91000
  Duke University
  Durham, NC 27708-1000
  Tel: 919 613 8159
  Fax: 919 613 8177
  Email: chauser@duke.edu
FEATURES
  Location/Qualifiers
  source
    /organism="Chlamydomonas reinhardtii"

```

```
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress condition I,
normalized, lambda zap II"
/Note="vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,
1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr
, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was
purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into lambda zap II
(Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
zap clones by superinfection with ExAssist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT 157 a 282 c 244 g 102 t 1 others
ORIGIN
Query Match 23.9%; Score 508.6; DB 12; Length 786;
Best Local Similarity 95.3%; Pred. No. 5e-78;
Matches 546; Conservative 0; Mismatches 24; Indels 3; Gaps 2;
QY 1554 CGCCCTGGCGCCACCGCTGGCGTGCACGAGGAGTGTTCGGGGGGGGCGGCTACCCCGA 1613
DB 95 CCCCCGGGCTGCAGGAATTCGGACACGAGGGGAGGTTTTCGGGGGGCGGCTACCCCGA 154
QY 1614 GATGTGGGCACTGATCAGCAGCAGACCTGCTCTGGTCCAGCCCGCCGAGAGTGGGA 1673
DB 155 GATGGTGGGCACTGATCAGCAGCAGACCTGCTCTGGTCCAGCCCGCCGAGAGTGGGA 214
QY 1674 GGCCCTGCTGGAGAGTGTGTACGCAAGGGCGGCTGGCCACCGCCCAAGAGGAGGA 1733
DB 215 GGGCTCTGCTGGAGAGTGTGTACGCAAGGGCGGCTGGCCACCGCCCAAGAGGAGGA 274
QY 1734 GATCAAGTGCCTGTCGCCGAGAGATCCCGGGGAGCTGCCCGCGGCTGTCTACGCCCC 1793
DB 275 GATCAAGTGCCTGTCGCCGAGAGATCCCGGGGAGCTGCCCGCGGCTGTCTACGCCCC 334
QY 1794 CAACACCTGAAGCCGCTGTCGCCCTCCGTTGGAGGGCAAGCGGCGCCGCGCCCAAGGT 1853
DB 335 CAACACCTGAAGCCGCTGTCGCCCTCCGTTGGAGGGCAAGCGGCGCCGCGCCCAAGGT 394
QY 1854 CGCACCAACCGCCCGCCATGGCGGCTGGCGGAGCACACCCCTTCGGGCGCCCTCGCC 1913
DB 395 CGCACCAACCGCCCGCCATGGCGGCTGGCGGAGCACACCCCTTCGGGCGCCCTCGCC 454
QY 1914 CGCGCGGCGCACCCCGCAAGGTGACCACTACAGCCCGCGGCTGCCCGCCACCGCCCAAGCC 1973
DB 455 CGCGCGGCGCACCCCGCAAGGTGACCACTACAGCCCGCGGCTGCCCGCCACCGCCCAAGCC 514
QY 1974 CAAGACCGCTGGCTCAAGTGGCGGTGAGGCTCCACACCTTCGACCTCGGAGACGG 2033
DB 515 CAAGACCGCTGGCTCAAGTGGCGGTGAGGCTCCACACCTTCGACCTCGGAGACGG 574
QY 2034 CGCTGCCTCCACGGCAACGGCAACGGTGGCTCGGCTCCAAAGACCTTCGGCTGCCAAGCC 2093
DB 575 CGCTGCCTCCACGGC-ACGGCAACGGTGGCTCGGCTTCAGAGACCTTCGGCTGCCAAGCC 633
QY 2094 CCTGGT--CTCGCGCGGCCACCGCAAGTCCGCC 2124
DB 634 CCTGGTCTTCGGGGGCCACCGGAAGTTCGCC 666
RESULT 2
AV394077 511 bp mRNA linear EST 23-APR-2002
LOCUS AV394077 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
DEFINITION cDNA clone CL18c08_r 5', mRNA sequence.
ACCESSION AV394077
VERSION AV394077.1 GI:6548293
KEYWORDS EST.
```

Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
1 (bases 1 to 511)
Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
A large scale structural analysis of cDNAs in a unicellular green
alga, Chlamydomonas reinhardtii. 1. Generation of 3433
non-redundant expressed sequence tags
DNA Res. 6 (6), 369-373 (1999)
20152988
MEDLINE
COMMENT
Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
Location/Qualifiers
1..511
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone_lib="CL18c08_r"
/clone_lib="Chlamydomonas reinhardtii C9"
/dev_stage="photoautotrophic growth"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 80 a 170 c 159 g 101 t 1 others
ORIGIN
Query Match 23.7%; Score 503.6; DB 10; Length 511;
Best Local Similarity 99.0%; Pred. No. 3.4e-77;
Matches 506; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 64 TTCGGTGTCAAGAGAACCGCGAACAGCTGCTGCGTAGAGCTTGCTGCGTCGCGACGC 123
DB 1 TTCGGTGTCAAGAGAACCGCGAACAGCTGCTGCGTAGAGCTTGCTGCGTCGCGACGC 60
QY 124 AAGTCCACCTCGCGCTCGGCTGTTACTGTGCGACCTGCTGCGCGCTGCGACATC 183
DB 61 AAGTCCACCTCGCGCTCGGCTGTTACTGTGCGACCTGCTGCGCGCTGCGACATC 120
QY 184 GTAGTGTGCTGCTGAGGTGCGCCCTTGGTCCAAAGCGGCGGCTGGCGATGTGACT 243
DB 121 GTAGTGTGCTGCTGAGGTGCGCCCTTGGTCCAAAGCGGCTGCTGGCGATGTGACT 180
QY 244 GGTGGCTGCGCTATTGAGCTGTGCAAGCGGGCCACCGCGTCATGACCATTTGCCCTCGC 303
DB 181 GGTGGCTGCGCTATTGAGCTGTGCAAGCGGGCCACCGCGTCATGACCATTTGCCCTCGC 240
QY 304 TACGACCAAGTACGCTGAGCGCTGGACACCTCGGTGCGTGGACATCATGGCGAGAAG 363
DB 241 TACGACCAAGTACGCTGAGCGCTGGACACCTCGGTGCGTGGACATCATGGCGAGAAG 300
QY 364 GTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGTGGATTGACACCCCTGG 423
DB 301 GTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGTGGATTGACACCCCTGG 360
QY 424 TTCCTGGCCAAAGGCTTGGGGCAAGACCGGTGCCAAGCTGTACGGCCCCCGCTCGCGGCT 483
DB 361 TTCCTGGCCAAAGGCTTGGGGCAAGACCGGTGCCAAGCTGTACGGCCCCCGCTCGCGGCT 420
QY 484 GACTACCTGGACACCAAGCGCTTCCCTGCTTCTCAAGGGCGGCTATTGAGGCTGCC 543
DB 421 GACTACCTGGACACCAAGCGCTTCCCTGCTTCTCAAGGGCGGCTATTGAGGCTGCC 480
QY 544 CGCGTGTGCTTCCGCGCCCGCGGAGGACT 574
DB 481 CGCGTGTGCTTCCGCGCCCGCGGAGGACT 511
RESULT 3
B0814687 502 bp mRNA linear EST 01-AUG-2002
LOCUS B0814687

DEFINITION	1030045B10.y1 C. reinhardtii CC-1690, Deflagellation (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION	U0814687
VERSION	BQ814687.1
KEYWORDS	GI:22062228
SOURCE	EST.
ORGANISM	Chlamydomonas reinhardtii. Chlamydomonas reinhardtii. Chlamydomonas reinhardtii. Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadales; Chlamydomonas.
REFERENCE	1. (bases 1 to 502)
AUTHORS	Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermott, J. P., Shrager, J., Silflow, C. and Stern, D.
TITLE	Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1030
JOURNAL	Unpublished (2002)
COMMENT	Contact: Charles Hauser DCMB Box 91000 Duke University Durham, NC 27708-1000 Tel: 919 613 8159 Fax: 919 613 8177 Email: chauser@duke.edu.

	Chlamydomonas reinhardtii.
ORGANISM	Chlamydomonas reinhardtii Chlamydomonas reinhardtii Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadales; Chlamydomonads.
REFERENCE	1 (bases 1 to 473)
AUTHORS	Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K., Nakamura,Y. and Tabata,S. Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii DNA Res. 7 (5), 305-307 (2000) 20539644
TITLE	Contact: Erika Asamizu
JOURNAL	The First Laboratory for Plant Gene Research
MEDLINE	Kazusa DNA Research Institute
COMMENT	Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizuka@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/ , Location/Qualifiers 1 . . 473
FEATURES	source

BASE COUNT	79 a	200 c	164 g	59 t	
ORIGIN					
Query Match	22.4%;				Score 476.6; DB 14;
Best Local Similarity	98.2%;				Length 502;
Matches	482; Conservative 0;				Mismatches 9; Indels 0; Gaps 0;

QY	1524	CCCGCAACAGCTGGACGAGGCTGACCGCAGCGCCTGGCGGCCACCGCTGGCGCGCGTGCAC	1583
Db	12	CCTCGTGC CGCTGGACGAGGCTGACCGCAGCGCCTGGCGGCCACCGCTGGCGCGTGCAC	71
QY	1584	CGAGTGTTTGGCGGCGCCGCTACCCGAGATGGTGCCCACTGCATCAGCCAGGACCT	1643
Db	72	CGAGTGTTTGGCGGCGCCGCTACCCGAGATGGTGCCCACTGCATCAGCCAGGACCT	131
QY	1644	GTCTCTGTCCAAAGCCGCCCAAGTGGAGGGGCTCTGGAGAGGTGTGTACGGCAA	1703
Db	132	GTCTCTGTCCAAAGCCGCCCAAGTGGAGGGGCTCTGGAGAGGTGTGTGTACGGCAA	191
QY	1704	GGCGCGCTGGCCACC GCCCAAGAGGAGGAGATCAAGGTGCCGCTTGC CGCAGAAAGATCCC	1763
Db	192	GGCGCGCTGGCCACC GCCCAAGAGGAGGAGATCAAGGTGCCGCTTGC CGCAGAAAGATCCC	251
QY	1764	CGCGCACCTGCCCGCTGTCTTACGCCCCCAACACCCTGAAGCCCGTGTCCGCGCTCCGT	1823
Db	252	CGCGCACCTGCCCGCGTGTCTTACGCCCCCAACACCCTGAAGCCCGTGTTCGCCCTCCGT	311
QY	1824	GGAGGGCAACGGCGCGCGGCCCAAGTGTGGCAACACCGCCCCGCCATGGGGCGGTG	1883
Db	312	GGAGGGCAACGGCGCGCGGCCCAAGTGTGGCAACACCGCCCCGCCATGGGGCGGTG	371

BASE COUNT	76 a	162 c	146 g	89 t	oxide ^a
ORIGIN					
Query Match	22.1%	Score 469.8;	DB 10;	Length 473;	
Best Local Similarity	99.6%	Pred. No. 2.3e-71;			
Matches 471;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
QY	188	TGTTGCTGCTGAGTGC	CCCCCTGTGTCC	AAGACGGCGCGCTTGGGCGATGTGACTGGTG	247
Db	1	TGTTGCTGCTGAGTGC	CCCCCTGTGTCC	AAGACGGCGCGCTTGGGCGATGTGACTGGTG	60
QY	248	GCCTGCCTATTGAGCT	TGTTCAACGGCGGCC	CAACCGGTCATGACCAITGCCCTCGCTACG	307
Db	61	GCCTGCCTATTGAGCT	TGTTCAACGGCGGCC	CAACCGGTCATGACCAITGCCCTCGCTACG	120
QY	308	ACCAGTACGCTGACGCT	GGGACACCTCGTNGT	CGTGTGGACATCATGGGGGAGAAGTCC	367
Db	121	ACCAGTACGCTGACGCT	GGGACACCTCGTNGT	CGTGTGGACATCATGGGGGAGAAGTCC	180
QY	368	GCTACTTCCACTCCAT	CAAGAGGGCGTGCAC	CGCGTGTGGATTGACCAACCCCTGGTTC	427
Db	181	GCTACTTCCACTCCAT	CAAGAGGGCGTGCAC	CGCGTGTGGATTGACCAACCCCTGGTTC	240
QY	428	TGSCCAAGGTC	TGGGCGACAGCGGCT	TCCAGCTGTACGCGCCCCGCTCGGCGCTGACT	487

```

Db 241 TGGCCAAAGTCTGGGGCAAGACGGCTCCAAAGCTGTACGGCCCCCGCTCCGGCGCTGACT 300
QY 488 ACCTGGAACACAAAGCGCTTGGCCCTGTTCTGCAAGCCCGCTATTAGGCTGCCGGG 547
Db 301 ACCTGGACAACACAAAGCGCTTGGCCCTGTTCTGCAAGCCCGCTATTAGGCTGCCGGG 360
QY 548 TGCTGCGCTTTCGGCCCGCGGAGGACTGGCTTCTGTTGCGCCCAAGCACTGGCACTCCGGCC 607
Db 361 TCGTGGCTTCGGCCCGCGGAGGACTGGCTTCTGTTGCGCCCAAGCACTGGCACTCCGGCC 420
QY 608 TGGTGGCGCTCTGCTGAAGAGCGAGTACCAAGCCCAAGGCGCGAGTTCACCAAG 660
Db 421 TGGTGGCGCTCTGCTGAAGAGCGAGTACCAAGCCCAAGGCGCGAGTTCACCAAG 473

RESULT 5
AV643281 506 bp mRNA linear EST 15-DEC-2000
LOCUS AV643281 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
DEFINITION cDNA clone HCL066h09_r 5', mRNA sequence.
ACCESSION AV643281
VERSION AV643281.1 GI:10786609
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
AUTHORS McDermott,J.P., Silflow,C., Stern,B. and Surzycki,R.,
TITLE Analyses of the Chlamydomonas reinhardtii genome: A Model,
JOURNAL Unicellular System for Analyzing Gene Function and Regulation in
MEDLINE Vascular Plants: project phase 2
COMMENT Unpublished (2000)
Contact: Elizabeth H. Harris
Duke University
Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlany@duke.edu
FEATURES
source
1. 506
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="HCL066h09_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT 77 a 170 c 160 g 99 t
ORIGIN
Query Match 22.0%; Score 466.8; DB 10; Length 506;
Best Local Similarity 99.6%; Pred. No. 7.5e-71;
Matches 468; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCTGTGCTTACAGCGCCGCCAGCGCGCTGCTATCTGTCATCAATGCCGCG 60
Db 37 ATGGCTGTGCTTACAGCGCCGCCAGCGCGCTGCTATCTGTCATCAATGCCGCG 96
QY 61 TCGTTCGGTGTCAAGAACCGCGAACAGCTGCTGCTGAGCTTGCTGCTGGTCCGCA 120
Db 97 TCGTTCGGTGTCAAGAACCGCGAACAGCTGCTGCTGAGCTTGCTGCTGGTCCGCA 156
QY 121 CGCAAGTCCACCTCGCGCTGGCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 157 CGCAAGTCCACCTCGCGCTGGCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 216
QY 181 ATCGTGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 217 ATCGTGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 276

```

```

QY 241 ACTGTGGCGCTGCCTATTGAGCTGTGTCACGGCGGCGCACCGCGTTCATGACCATTTGCCCT 300
Db 277 ACTGTGGCGCTGCCTATTGAGCTGTGTCACGGCGGCGCACCGCGTTCATGACCATTTGCCCT 336
QY 301 CGCTAGCACCACTGACCTGACCGCTGGGACACCTCGGTGCTGTCGACATCATGGCGAG 360
Db 337 CGCTAGCACCACTGACCTGACCGCTGGGACACCTCGGTGCTGTCGACATCATGGCGAG 396
QY 361 AAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGTGGATTGACACCCC 420
Db 397 AAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGTGGATTGACACCCC 456
QY 421 TGGTTCCTGGCAAGGTCTGGGCAAGACCGGTCCTCAAGCTGTACGCCCC 470
Db 457 TGGTTCCTGGCAAGGTCTGGGCAAGACCGGTCCTCAAGCTGTACGCCCC 506

RESULT 6
AW757933 475 bp mRNA linear EST 03-MAY-2000
LOCUS 874004E12.y1 C. reinhardtii CC-1690, Lambda Zap II Chlamydomonas
DEFINITION reinhardtii cDNA, mRNA sequence.
ACCESSION AW757933
VERSION AW757933.1 GI:7687285
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
AUTHORS McDermott,J.P., Silflow,C., Stern,B. and Surzycki,R.,
TITLE Analyses of the Chlamydomonas reinhardtii genome: A Model,
JOURNAL Unicellular System for Analyzing Gene Function and Regulation in
MEDLINE Vascular Plants: project phase 2
COMMENT Unpublished (2000)
Contact: Elizabeth H. Harris
Duke University
Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlany@duke.edu
FEATURES
source
1. 475
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP (acetate-containing) medium in
light, TAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with ExAssist (Stratagene)
phage."
BASE COUNT 74 a 160 c 143 g 98 t
ORIGIN
Query Match 21.6%; Score 458; DB 10; Length 475;
Best Local Similarity 97.9%; Pred. No. 2.4e-69;
Matches 464; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 77 AGACGGCAACACAGCTGCTGCTGAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 136
Db 2 AGACGGCAACACAGCTGCTGCTGAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 61
QY 137 GCTCGGCTGTTACTGCTGCCACTGCTGCTGCCACTTGGCGCGCTGCACATCATGATGTTGCTG 196

```



```

|||||
Db 62 GCTGGCTGTACTGGTCCCACTGGTGCACCTTGTCTCGTGACATCTGTGAIGTTGCTG 121
QY 197 CTGAGGTGCGCCCTTGTTCAGAGAGGGCGGCTGGGGATGTGACTGGTGGCCCTGCCTA 256
Db 122 CTGAGGTGCGCCCTTGTTCAGAGAGGGCGGCTGGGGATGTGACTGGTGGCCCTGCCTA 181
QY 257 TTGAGCTGTCAAGCGGGCGGCGGCTGATGACATGACATGACATGACATGACATGAC 316
Db 182 TTGAGCTGTCAAGCGGGCGGCGGCTGATGACATGACATGACATGACATGACATGAC 241
QY 317 CTGAGCGCTGGGACACCTCGGCTGTGCTGGACATCATGGGCGAGAGGTGCGCTACTTCC 376
Db 242 CTGAGCGCTGGGACACCTCGGCTGTGCTGGACATCATGGGCGAGAGGTGCGCTACTTCC 301
QY 377 ACTCCATCAAGAGGGCGGCGGCTGATGACATGACATGACATGACATGACATGACATGAC 436
Db 302 ACTCCATCAAGAGGGCGGCGGCTGATGACATGACATGACATGACATGACATGACATGAC 361
QY 437 TCTGGGCAAGACCGGCTCCCAAGCTGTACGCGCCCGGCTCGGGGCTGACTTACCTGGACA 496
Db 362 TCTGGGCTAGACCGGCTCCCAAGCTGTACGCGCCCGGCTCGGGGCTGACTTACCTGGACA 421
QY 497 ACCACAGCGCTTGGCCCTCTTCTGCAAGCGCGCTATTGAAGCTGCCCGGTGC 550
Db 422 ACCACAGCGCTTGGCCCTCTTCTGCAAGCGCGCTATTGAAGCTGCCCGGTGC 475

RESULT 7
LOCUS AV629488 461 bp mRNA linear EST 15-DEC-2000
DEFINITION AV629488 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas
reinhartdii cDNA clone LCL059f09_r 5', mRNA sequence.
ACCESSION AV629488
VERSION AV629488.1 GI:10792122
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
REFERENCE 1 (bases 1 to 461)
AUTHORS Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..461
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="LCL059f09_r"
/clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO2"
note="Vector: pBluescriptII SK-; Site.1: EcoRI; Site.2:
XhoI; The cDNA library was constructed from cells cultured
in a carbon stress acclimatized condition in which carbon
dioxide concentration in the bubbling gas was changed from
5% to 0.04%."
BASE COUNT 73 a 156 c 143 g 89 t
ORIGIN
|||||

Query Match 21.6%; Score 457.8; DB 10; Length 461;
Best Local Similarity 99.6%; Pred. No. 2.6e-69;
Matches 459; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 179 ACATCGTATGTTGCTGCTAGGTGCGCCCTTGTCTCAAGAGCGGCGCTGGCGGATG 238
|||||

```

```

Db 1 ACATCGTATGTTGCTGCTAGGTGCGCCCTTGTCTCAAGAGCGGCGCTGGCGGATG 60
QY 239 TGACTTGTGGCTGCTGCTATTTGAGCTGGTCAAGCGGCGGCGGCGCTGATGACCATGGCC 298
Db 61 TGACTTGTGGCTGCTGCTATTTGAGCTGGTCAAGCGGCGGCGGCGCTGATGACCATGGCC 120
QY 299 CTCGCTAGACACAGTACGCTGACGCTGGGACACCTCGGCTGGTCTGGACATCATGGGCG 358
Db 121 CTCGCTAGACACAGTACGCTGACGCTGGGACACCTCGGCTGGTCTGGACATCATGGGCG 180
QY 359 AGAAGTCCGCTACTTCCACTCCATCAAGAGGGCGGCTGCACCGGCTGTGGATTGACAC 418
Db 181 AGAAGTCCGCTACTTCCACTCCATCAAGAGGGCGGCTGCACCGGCTGTGGATTGACAC 240
QY 419 CCTGGTTCCTGCGCCAAAGTCTGGGCAAGACGCGCTCCCAAGCTGTACGCGCCCGCTCG 478
Db 241 CCTGGTTCCTGCGCCAAAGTCTGGGCAAGACGCGCTCCCAAGCTGTACGCGCCCGCTCG 300
QY 479 GCGTGTACTGCTGACCAACCAAGCGCTTCGCGCCCTTCTGCAAGCGGCTATTGAGG 538
Db 301 GCGTGTACTGCTGACCAACCAAGCGCTTCGCGCCCTTCTGCAAGCGGCTATTGAGG 360
QY 539 CTGCCCCGGTGTGCGCTTCGCGCCCGGCGGAGGACTGCGCTTTCGTGGCCCAAGGACTGG 598
Db 361 CTGCCCCGGTGTGCGCTTCGCGCCCGGCGGAGGACTGCGCTTTCGTGGCCCAAGGACTGG 420
QY 599 ACTCGGCGCTGCTGCGCGCTGCTGCTGAAGGACGAGTACCAG 639
Db 421 ACTCGGCGCTGCTGCGCGCTGCTGCTGAAGGACGAGTACCAG 461

RESULT 8
LOCUS AY109531
DEFINITION Zea mays CLL198_1 mRNA sequence.
ACCESSION AY109531
VERSION AY109531.1 GI:21213285
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2147)
AUTHORS Rainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overigo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 2147)
AUTHORS Coe,E.C.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA

FEATURES
source
1..2147
/organism="Zea mays"
/db_xref="MaizeDB:630100"
/db_xref="taxon:4577"
/clone="CCL198_1"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project."
BASE COUNT 400 a 582 c 604 g 365 t 196 others
ORIGIN

Query Match 21.5%; Score 456.8; DB 11; Length 2147;
Best Local Similarity 58.5%; Pred. No. 4.7e-69;

```

	Matches	917;	Conservative	0;	Mismatches	555;	Indels	96;	Gaps	6
QY	161	GTCCCATCTGCGCGCTGGACATCGT	GTGCTGAGTGTGCTGAGTGC	CGCCCTTGGTCCAA	GA	220				
Db	215	GCSCAGCCCGCATGAACGTCGTCT	TCGTCGSCCGGAGATGCGCGTGGAGCA	GA	274					
QY	221	CGGCGCGCTTGGCGATGTGACTGT	GGTCCCTATTTAGCTGTTC	AAGCGCGCCACC	280					
Db	275	CCGCGCGCTTCGCGCGACGNNNN	NNNNNNNNNNNNNNNNNNNN	NATGGCCGCAACCGGCACC	334					
QY	281	GCCTCATGACATTCGCCCTCGCT	TACGACCATGACGTGACGCT	TGGGACACCTTCGGTGG	340					
Db	335	GTCTATGTGCTCTCTCCCGCT	TACGACCATGACGTGACGCT	TGGGACACCTTCGGTGG	394					
QY	341	TCGTGGACATCA	-----TGGCGGAGAGGTCCGCT	TACTTCCATCCATCAAGA	388					
Db	395	TGTCGGAGATCAAGATGGGAGAC	GGGTACGAGACGGTTCAGGTTC	TTCACCTGCTACAAGC	454					
QY	389	AGGCGTGACCGGCTGTGGATT	CACACCCCTGGTCTCTGCGCAAG	GTCTGGGCAAGA	448					
Db	455	CGGAGTGGACCGGCTGTCTG	TGACCAACCACATGTTCT	TGGAGAGGNNNNNNNNNN	514					
QY	449	CCGCTCTCAAGCTTACGGCC	CCCGCTCCGCGCTGACT	TCTGTCGACCAACACAGCGCT	508					
Db	515	NNNNNNNNNNNNCTACGGCT	CTGCTGTAACGGACTACAGG	GACCAACAGCTTCGCT	574					
QY	509	TCGCCCTGTTCTGAAGCCGCT	TATTAGGCTGCCCGTCTGCC	TTCGCTTCGCTTCGCT	561					
Db	575	TCAGCTGCTATGCGGACGAC	TTGAAGTCCAAAGGATTC	TGAGCTTCAACCAACAC	634					
QY	562	-----CCCGCGGAGGACT	CGCTTCGTCGCGCAACGACT	GCACCTCG	604					
Db	635	CATACTTCTCGGACCATAC	GGGAGGACGCTGCTGCT	TGCAACGACTGGCACCG	694					
QY	605	CCCTGTGTCGCGTCTGCT	GAAGGACGAGTACCACCC	AAGGGCCAGTTCACCAAGGCA	664					
Db	695	GCCCTCTCTCGTGTACCT	CAAGACCACTACCACTCC	ACCGGCAATCACAGGAGCGCA	754					
QY	665	AGTCGGTCTGGCTATCCACA	CACTCGCTTCCAGGGCGCAT	GTGGGAGGAGGCTTCA	724					
Db	755	AGACGCTTCTGATCCACA	CACTCTCTACCAAGGCGG	GTGCTTCCTCCGACTACC	814					
QY	725	AGGACAGAACTGCCCGAG	CGCGCTTTGACAAGCTGG	CGCTTCGAGCGCTATGCCA	784					
Db	815	CGGAGCTGAACCTCCCG	GAGAGATTCAAGTCTGCT	TTCGATTCATCGACGGCTACGAG	874					
QY	785	AGTTTACATTGAGGCC	ACCCCATGAGGAGGACGAG	AAGCCCGCTGACGGGAA	844					
Db	875	AG-----	-----CCGCTGGAAG	-----	887					
QY	845	CCTACAAGAAGATCAACT	GGCTGAAGGTTGCAATTAT	CGCCCGCAACAGCTGGT	904					
Db	888	---CCGGAAGATCAACT	GGATGAAGCGCGGATCCT	CGAGCGCGACAGGCTCCTCACCG	943					
QY	905	TGTGCCCAACTACGCG	ACGAGATCGTCCGAT	CGCCGCGCTGTGGAGCTGGACA	964					
Db	944	TCAGCCCCCTACTAC	CGGAGGACTATCTCCG	GCATCGCCAGGGCTCGGAGCTCGACA	1003					
QY	965	CCGTATCCGGGCCAAG	GCGCATGTGAACGCGCAT	TGAGCACTTGAGCACTGGA	1024					
Db	1004	ACATCATGGCTCAC	CGGSCATACCGGCTG	TCAACGSCATGAGCGTCA	1063					
QY	1025	ACCCCAAGACGCAAG	TCTGCTGCGCCCTAC	ACACAGAACAGCGTCTACGCGGCA	1084					
Db	1064	ACCCACAGGAGCAAG	TACATCGCGCTGAAGT	TACAGCTGTGACGCGCGTGGAGGCCA	1123					
QY	1085	AGGCGCGCCCAAG	GAGGCCCTCGAGCGCG	AGTGGGCTTCCTGTGGACCCACCGCC	1144					
Db	1124	AGGCGCTGAACAAG	GAGGCGCTCGAGCGGAG	TGCGGCTCCCGGTGGACCGGAACATCC	1183					
QY	1145	CCCTGTTCGCTTCA	TCGCGCGCTGGAGGAC	GAGAGGAGTGTGGACATCATCTG	1204					
Db	1184	CGCTGTGCGCTTCA	TCGCGAGCTGGAAG	AGCAGAGGCGCCGACGCTATGCGCGCC	1243					

QY	1205	CCTGCCACAGATCTTGGCCACCCCAAG---GTCCAGATCGCATCTCTGGTACCGGCA	1261
Db	1244	CCATCCCGCAGCTCAATGGAGATGGTGGAGAGCTGCATCGTCTGCTGGGACCGGGCA <td>1303</td>	1303
QY	1262	AGCGCCCTACGAGAAGCTGGTGAAGCCATCGGCACCAAGTACAAGGGCGCGGCCAAGG <td>1321</td>	1321
Db	1304	AGAGAAGTTTCGAGCGCATGCTCATGAGCGCGAGGAGAAGTCCAGGCAAGGTGCGCG <td>1363</td>	1363
QY	1322	GCCTGGTCAAGTTCTCGGCGCCCTTGGCGCACATGCTCACCGCGCGCGCGACTTCATGC <td>1381</td>	1381
Db	1364	CGTGGTCAATTCACGCGCGCTGGCGCACCATCATGCGCGCGCGCGACGTGCTCG <td>1423</td>	1423
QY	1382	TGGTGGCCCTTCGCGCTTCGAGCCCTCGGGCGCTGATCCAGCTGCAGCCCATGCACCTACG <td>1441</td>	1441
Db	1424	CCGTACACGCGCTTCGAGCCCTCGGGCGCTCATCCAGCTGCAGGGATCGGATACCGAA <td>1483</td>	1483
QY	1442	CCGTGCCCTGGTAGCTCCACCGGGCGCTGGTTCGACACCGTCAAGGAGGCGCTCACCG <td>1501</td>	1501
Db	1484	CGCCCTCGCGCTGGGGTCCACCGGTGAGCTGCTGCACCATCATCGAAGCAAGACCG <td>1543</td>	1543
QY	1502	GCCTCCACATGGCGCCCTGA-----ACCCGACAAAGCTGGACGAGCTGACGCCGAGC <td>1555</td>	1555
Db	1544	GGTTCACATGGCGCGCTCAGCGTGCAGTGAACGCTGTGGAGCGCGCGACGCTCAAGA <td>1603</td>	1603
QY	1556	CCCTGGCGCGCAGCTGGCGCGCTGCAGCGAGGTGTTTCGGGCGCGCTACCCCGAGA <td>1615</td>	1615
Db	1604	AGTGGGCGACCACTTTCGACGCGCCATCAAGGTGTGCGGACGCCGCGCTACGAGGAGA <td>1663</td>	1663
QY	1616	TGGTGGCAACTGCATCAGCAGCAGCACTGTCTGTGTCGAAGCGCGCCAGAGTGGGAGG <td>1675</td>	1675
Db	1664	TGGTGAAGACTGCATCATCCAGATCTCTCTGGAGGCGCTTCCAGAACTGGGAGA <td>1723</td>	1723
QY	1676	GCCTGCTG 1683 <td></td>	
Db	1724	ACGTGCTG 1731 <td></td>	
RESULT 9			
BM003295			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

```

1. .666
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 2lgr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized
), Lambda zap II"
/notes=vector: pBluescript II SK-; Site1: EcoRI; Site2:
XhoI; Stress condition II library, constructed by John

```

Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH₄⁺ - containing) and shifted to TAP - NO₃⁻ (24hrs); H2 production conditions (0, 12hr, 24hr) see Mellis et al., (2000) Plant Phys. 122: 127-135; TAP + H2O₂ (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

BASE COUNT 121 a 244 c 224 g 77 t
ORIGIN

Query Match 21.5%; Score 456; DB 13; Length 666;
Best Local Similarity 100.0%; Pred. No. 5.6e-69;

Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1669 TGGAGGCGCTGTGAGAGGTGTGTACGGCAAGGGGGCGTGGCCACCGCAAGAAG 1728

Db 1 TGGAGGGCGCTGTGAGAGGTGTGTACGGCAAGGGGGCGTGGCCACCGCAAGAAG 60

QY 1729 GAGGAGATCAAGTGGCGCTTGCAGAGATCCCGGGGACCTGCCCGCGTGTCTTAC 1788

Db 61 GAGGAGATCAAGTGGCGCTTGCAGAGATCCCGGGGACCTGCCCGCGTGTCTTAC 120

QY 1789 GCGCCCAACACCTGAAGCGCGTGTCCGCTCCGTGGAGGGCAACGGCGCGCGCGCC 1848

Db 121 GCGCCCAACACCTGAAGCGCGTGTCCGCTCCGTGGAGGGCAACGGCGCGCGCGCC 180

QY 1849 AAGGTGGGACCAACCGCCCGCCATGGCGGTGGCGGCGGACACCGCCCTCGGGCGCC 1908

Db 181 AAGGTGGGACCAACCGCCCGCCATGGCGGTGGCGGCGGACACCGCCCTCGGGCGCC 240

QY 1909 TCGCCCGCGCGCCACCGCCCAAGGTGACACCTACAAGCGCGCCCTGCCCGCCACCGCC 1968

Db 241 TCGCCCGCGCGCCACCGCCCAAGGTGACACCTACAAGCGCGCCCTGCCCGCCACCGCC 300

QY 1969 AAGCCCAAGCGCTGGCGCTCAAGCTGGCGGTGAGGCTCCACCACTCGACCTCGGAG 2028

Db 301 AAGCCCAAGCGCTGGCGCTCAAGCTGGCGGTGAGGCTCCACCACTCGACCTCGGAG 360

QY 2029 AAGCGCGCTGCTCCAAAGCGACGCGACGTGCTCGGCTCCAAAGACCTCGGCTGCC 2088

Db 361 AAGCGCGCTGCTCCAAAGCGACGCGACGTGCTCGGCTCCAAAGACCTCGGCTGCC 420

QY 2089 AAGCCCGCTGCTCGCGCGCCACCGCAAGTCCGCC 2124

Db 421 AAGCCCGCTGCTCGCGCGCCACCGCAAGTCCGCC 456

RESULT 10

AV622787

LOCUS

AV622787 458 bp mRNA linear EST 15-DEC-2000

DEFINITION

reinhardtii cDNA clone LC054e04_r 5', mRNA sequence.

ACCESSION

AV622787

VERSION

AV622787.1 GI:10771964

KEYWORDS

EST

SOURCE

Chlamydomonas reinhardtii.

ORGANISM

Chlamydomonas reinhardtii.

REFERENCE

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

Chlamydomonadales; Chlamydomonas.

Asamizu, E., Miura, K., Kuchino, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,

Nakamura, Y. and Tabata, S.

Generation of expressed sequence tags from low-CO2 and high-CO2

adapted cells of Chlamydomonas reinhardtii

DNA Res. 7 (5), 305-307 (2000)

JOURNAL

MEDLINE

20539644

COMMENT

Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizukazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

source

Location/Qualifiers

1..458

/organism="Chlamydomonas reinhardtii"

/strain="C9"

/db_xref="taxon:3055"

/clone="LC054e04_r"

/notes="Vector: pBluescriptII SK-; Site.1: EcoRI; Site.2:

XhoI; The cDNA library was constructed from cells cultured

in a carbon stress acclimatized condition in which carbon

dioxide concentration in the bubbling gas was changed from

5% to 0.04%"

5% to 0.04%"

BASE COUNT 91 a 147 c 146 g 74 t

ORIGIN

Query Match 20.9%; Score 444.4; DB 10; Length 458;

Best Local Similarity 99.6%; Pred. No. 5.3e-67;

Matches 456; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 536 AGGTGCGCGCTGTGCTCCCTTGGCGCCCGCGAGGACTCGTCTTCGTGGCCCAACGACT 595

Db 1 AGGTGCGCGCTGTGCTCCCTTGGCGCCCGCGAGGACTCGTCTTCGTGGCCCAACGACT 60

QY 596 GGCACCTCCGCGCTGTGCTGCTGAAGACAGTACCAGCCCAAGGGCCAGTTCA 655

Db 61 GGCACCTCCGCGCTGTGCTGCTGAAGACAGTACCAGCCCAAGGGCCAGTTCA 120

QY 656 CCAAGGCCAAGTGGTGGCTATC-CACAACTCCCTTCCAGGCGCCCATGTGGAG 714

Db 121 CCAAGGCCAAGTGGTGGCTATC-CACAACTCCCTTCCAGGCGCCCATGTGGAG 180

QY 715 GAGGCTTTCAGGACAGAGCTGCCCGCCGCTTTGACAAGCTGGCTTCTCGGAC 774

Db 181 GAGGCTTTCAGGACAGAGCTGCCCGCCGCTTTGACAAGCTGGCTTCTCGGAC 240

QY 775 GGTATGCCAAGGTTTACACTGAGGCGCCACCCCATGGAGGAGACGAGAGCCCCGCTG 834

Db 241 GGTATGCCAAGGTTTACACTGAGGCGCCACCCCATGGAGGAGACGAGAGCCCCGCTG 900

QY 835 ACGGGAAGACCTTACAAAGATCAACTGGCTGAAGGTGGCATTTATCGCGCGGACAAG 894

Db 301 ACGGGAAGACCTTACAAAGATCAACTGGCTGAAGGTGGCATTTATCGCGCGGACAAG 360

QY 895 CTGGTCACTGTGCGCCCAACTACGCGACCGAGATCGCTGCCGATCCCGCGGCGGTG 954

Db 361 CTGGTCACTGTGCGCCCAACTACGCGACCGAGATCGCTGCCGATCCCGCGGCGGTG 420

QY 955 GAGCTGCACACCGTCATCCGCGCCCAAGGGCATTGAGG 992

Db 421 GAGCTGCACACCGTCATCCGCGCCCAAGGGCATTGAGG 458

RESULT 11

AV628313

LOCUS

AV628313 560 bp mRNA linear EST 15-DEC-2000

DEFINITION

reinhardtii cDNA clone LCL039h11_r 5', mRNA sequence.

ACCESSION

AV628313

VERSION

AV628313.1 GI:10790947

KEYWORDS

EST

SOURCE

Chlamydomonas reinhardtii.

ORGANISM

Chlamydomonas reinhardtii.

REFERENCE

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

Chlamydomonadales; Chlamydomonas.

1 (bases 1 to 560)

Asamizu, E., Miura, K., Kuchino, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,

Nakamura, Y. and Tabata, S.

Generation of expressed sequence tags from low-CO2 and high-CO2

adapted cells of Chlamydomonas reinhardtii

DNA Res. 7 (5), 305-307 (2000)

JOURNAL

MEDLINE

20539644

JOURNAL adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

source
1..560
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="LCL039h11_r"
/clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO₂"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a carbon stress acclimatized condition in which carbon
dioxide concentration in the bubbling gas was changed from
5% to 0.04%"

BASE COUNT 92 a 186 c 173 g 109 t
ORIGIN

Query Match 20.9%; Score 443.8; DB 10; Length 560;
Best Local Similarity 99.6%; Pred. NO. 6.9e-67;
Matches 445; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCTGTGCTCTACACAGCGCCGCCAGCAGCGCGCTCTATCGTCATCAATGCCGCG 60
Db 114 ATGGCTGTGCTCTACACAGCGCCGCCAGCAGCGCGCTCTATCGTCATCAATGCCGCG 173
QY 61 TCGTTCCGGTCAAGAACGCGCAACAGCTGCTCGTGAGCTTGCTCGTGCGCTCCGCA 120
Db 174 TCGTTCCGGTCAAGAACGCGCAACAGCTGCTCGTGAGCTTGCTCGTGCGCTCCGCA 233
QY 121 CGCAAGTCCACCTCGCGCTCGGCTGTACTGTGCGCACTGGTGCGCACTGGCGCTGGAC 180
Db 234 CGCAAGTCCACCTCGCGCTCGGCTGTACTGTGCGCACTGGTGCGCACTGGCGCTGGAC 293
QY 181 ATCGTATGTTGCTCTGAGGTCGCGCTTGTGTCGAAGAGCGCGCGCTCGGCGATGTG 240
Db 294 ATCGTATGTTGCTCTGAGGTCGCGCTTGTGTCGAAGAGCGCGCGCTCGGCGATGTG 353
QY 241 ACTGTGGCGCTGCTATTGAGCTGTGTCAGCGCGCGCACCGCTCATGACCATTTGCCCT 300
Db 354 ACTGTGGCGCTGCTATTGAGCTGTGTCAGCGCGCGCACCGCTCATGACCATTTGCCCT 413
QY 301 CGGTACGACAGTACGCTGACGCTGGGACACCTCGGTGCTGGACATCATGGGCGAG 360
Db 414 CGGTACGACAGTACGCTGACGCTGGGACACCTCGGTGCTGGACATCATGGGCGAG 473
QY 361 AAGGTCGGCTTCTTCCACTCCATCAAGAAGGCGGTGACCGCGTGTGGATTGACCAACC 420
Db 474 AAGGTCGGCTTCTTCCACTCCATCAAGAAGGCGGTGACCGCGTGTGGATTGACCAACC 533
QY 421 TGGTCTCGCCCAAGGCTTGGGCGAAG 447
Db 534 TGGTCTCGCCCAAGGCTTGGGCGAAG 560

RESULT 12
AV629117 436 bp mRNA linear EST 15-DEC-2000
LOCUS AV629117 Chlamydomonas reinhardtii 5% to 0.04% CO₂ Chlamydomonas
DEFINITION reinhardtii cDNA clone LCL052d05_r 5', mRNA sequence.
ACCESSION AV629117 GI:10791751
VERSION AV629117.1
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaeae; Chlamydomonas.
1 (bases 1 to 436)

AUTHORS

Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
Nakamura, Y. and Tabata, S.
Generation of expressed sequence tags from low-CO₂ and high-CO₂
adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

source
1..436
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="LCL052d05_r"
/clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO₂"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a carbon stress acclimatized condition in which carbon
dioxide concentration in the bubbling gas was changed from
5% to 0.04%"

BASE COUNT 66 a 145 c 138 g 87 t
ORIGIN

Query Match 20.4%; Score 432.8; DB 10; Length 436;
Best Local Similarity 99.5%; Pred. NO. 5.2e-65;
Matches 434; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 140 CGGCTGTACTGTGCGACTGTGCCACTTGCCTGCGCTGGACATCGATGCTGCTGCTG 199
Db 1 CGGCTGTACTGTGCGACTGTGCCACTTGCCTGCGCTGGACATCGATGCTGCTGCTG 60
QY 200 AGTGCGCCCTTGTGTCGAAGACGGCGCGCTGGCGGATGTGACTGTGGCTGCTTATG 259
Db 61 AGTGCGCCCTTGTGTCGAAGACGGCGGCTGGCGGATGTGACTGTGGCTGCTTATG 120
QY 260 AGTGCTGAAGCGCGCGCACCGCGCTCATGACCATTTGCCCTCGCTAGCACGATGCTG 319
Db 121 AGTGCTGAAGCGCGCGCACCGCGCTCATGACCATTTGCCCTCGCTAGCACGATGCTG 180
QY 320 ACGCCCTGGGACACTCGGTGCTGTGACATCATGGCGAGAGGTCGCTACTTCCACT 379
Db 181 ACGCCCTGGGACACTCGGTGCTGTGACATCATGGCGAGAGGTCGCTACTTCCACT 240
QY 380 CCATCAAGAAGGCGGTGCACCGCGTGTGGATTGACCAACCCCTGGTTCCTGGCCAAAGTCT 439
Db 241 CCATCAAGAAGGCGGTGCACCGCGTGTGGATTGACCAACCCCTGGTTCCTGGCCAAAGTCT 300
QY 440 GGGCAAGACCGGCTCCAAAGCTGTACGGCCCGCGCTCGGCGCTGACTTGGACAACC 499
Db 301 GGGCAAGACCGGCTCCAAAGCTGTACGGCCCGCGCTCGGCGCTGACTTGGACAACC 360
QY 500 ACAAGCGCTTCGCGCTTCTTGTCAAGCGCGCTATTGAGGCTGCCCGGCTGCGCCCTCG 559
Db 361 ACAAGCGCTTCGCGCTTCTTGTCAAGCGCGCTATTGAGGCTGCCCGGCTGCGCCCTCG 420
QY 560 GCCCGCGCGAGGACTG 575
Db 421 GCCCGCGCGAGGACTG 436

RESULT 13
BE024926 552 bp mRNA linear EST 06-JUN-2000
LOCUS BE024926 894006H08.v1 C. reinhardtii CC-1690, normalized, Lambda Zap II
DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BE024926 GI:8287367
VERSION BE024926.1
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Grossman, A., Davies, J., Federspiel, N., Harris, E., Lefebvre, P.,
McDermott, J. P., Sillfow, C., Stern, D. and Surzycki, R.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants; project phase 2
Unpublished (2000)
Contact: Elizabeth H. Harris
DCMB Box 91000
Duke University
Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlamy@duke.edu.

FEATURES
source

1. .552
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 2lgr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in YAP (acetate-containing) medium in the
light, YAP medium in the dark, HS (minimal) medium in
ambient levels of CO₂ and HS medium bubbled with 5% CO₂.
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with ExAssist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
91 a 185 c 167 g 109 t

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 19.9%; Score 422.8; DB 10; Length 552;
Matches 424; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCTGTGGCTCTACCAAGCGCGCCAGCGCGCTCTATCGTCAATCAATGCGCG 60
Db 127 ATGGCTGTGGCTCTACCAAGCGCGCCAGCGCGCTCTATCGTCAATCAATGCGCG 186
QY 61 TCGTTGGGTGTCAGAAAGCGCGAACAGCTGCTGCGTGAGCTTGTCTGCGCTCGCA 120
Db 187 TCGTTGGGTGTCAGAAAGCGCGAACAGCTGCTGCGTGAGCTTGTCTGCGCTCGCA 246
QY 121 CGCAAGTCCACCTCGCGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 247 CGCAAGTCCACCTCGCGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306
QY 181 ATCGGTGATGGTGTGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 307 ATCGGTGATGGTGTGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366
QY 241 ACTGTTGCGCTTGCCTATTGAGTGTGCTGAGCGCGCGCCACCGCGTCAATGCGCCCT 300
Db 367 ACTGTTGCGCTTGCCTATTGAGTGTGCTGAGCGCGCGCCACCGCGTCAATGCGCCCT 426
QY 301 CGCTACGACCACTAGCTGAGCGCTGGGACACCTCGGTGGTGGTGGTGGTGGTGGTGG 360
Db 427 CGCTACGACCACTAGCTGAGCGCTGGGACACCTCGGTGGTGGTGGTGGTGGTGGTGG 486
QY 361 AAGTCCGCTACTTCCACTCCATCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Db 487 AAGTCCGCTACTTCCACTCCATCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 546
QY 421 TGGTTC 426
|||||

Db 547 TGGTTC 552

RESULT 14
AV641583
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AV641583 517 bp mRNA linear EST 15-DEC-2000
AV641583 Chlamydomonas reinhardtii 5% CO₂ Chlamydomonas reinhardtii
CDNA clone HCL036g08_r 5', mRNA sequence.
AV641583
AV641583.1 GI:10784911
EST.
Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

1 (bases 1 to 517)
Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
Nakamura, Y. and Tabata, S.
Generation of expressed sequence tags from low-CO₂ and high-CO₂
adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
20539644
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source

1. .517
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="HCL036g08_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO₂"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 19.8%; Score 421.2; DB 10; Length 517;
Matches 423; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCTGTGGCTCTACCAAGCGCGCCAGCGCGCTCTATCGTCAATCAATGCGCG 60
Db 92 ATGGCTGTGGCTCTACCAAGCGCGCCAGCGCGCTCTATCGTCAATCAATGCGCG 151
QY 61 TCGTTGGGTGTCAGAAAGCGCGAACAGCTGCTGCGTGAGCTTGTCTGCGTCCGCA 120
Db 152 TCGTTGGGTGTCAGAAAGCGCGAACAGCTGCTGCGTGAGCTTGTCTGCGTCCGCA 211
QY 121 CGCAAGTCCACCTCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 212 CGCAAGTCCACCTCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 271
QY 181 ATCGGTGATGGTGTGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 272 ATCGGTGATGGTGTGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 331
QY 241 ACTGTTGCGCTTGCCTATTGAGTGTGCTGAGCGCGCGCCACCGCGTCAATGCGCCCT 300
Db 332 ACTGTTGCGCTTGCCTATTGAGTGTGCTGAGCGCGCGCCACCGCGTCAATGCGCCCT 391
QY 301 CGCTACGACCACTAGCTGAGCGCTGGGACACCTCGGTGGTGGTGGTGGTGGTGGTGG 360
Db 392 CGCTACGACCACTAGCTGAGCGCTGGGACACCTCGGTGGTGGTGGTGGTGGTGGTGGTGG 451
QY 361 AAGTCCGCTACTTCCACTCCATCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Db 452 AAGTCCGCTACTTCCACTCCATCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 511
|||||

APPLICATION NUMBER: US/08/941,445A
FILING DATE: 30-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,855
FILING DATE: 30-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Winner, Ellen P.
REGISTRATION NUMBER: 28,547
REFERENCE/DOCKET NUMBER: 89-97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2542 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: CDS
LOCATION: 453..2282
US-08-941-445A-6

Query Match 21.1%; Score 447.6; DB 3; Length 2542;
Best Local Similarity 58.5%; Pred. No. 4.5e-67;
Matches 932; Conservative 0; Mismatches 564; Indels 96; Gaps 5;

QY 171 CGCGTGGACATCGTGGTGGCTGCTGCTGAGTGGCGCCCTTGGTTCGACGAGCGGCGCCCT 230
DB 695 CGCGATGAACGTCGTGCTGGCGCCGAGATGGCCCTCGGAGCAAGACCGCGCCCT 754
QY 231 GGGCGATGTGACTGTGGTGGCTGCTATTGAGCTGGTCAAGCGGGCGCCGCTCATGAC 290
DB 755 CGGTGACGTCCTGGTGGCTGCCCTGCCATGGTGGCATGGGACAGGCGTCATGGT 814
QY 291 CATTTGCCCTTCGCTACGACACGATGACGCTGGGACACCTCGTGGTGGTGGACAT 350
DB 815 GATCTCTCTCGGTAGCAGCAGTACAGGAGCGCTGGGATACCGGTTGGCTGAGAT 874
QY 351 CATGG-----GGAGAGGTGCGCTACTTCCACTCCATCAAGAAGGGCGTGCA 398
DB 875 CAAAGGTGCAGACAGTACGAGAGGTGAGGTTTTCATTTGCTACAAAGCGTGGAGTGA 934
QY 399 CGCGGTGTGGATGACACCGCTGTTCTTGGCCAAAGGTCTGGGCAAGACCGSCTCCAA 458
DB 935 CCGTGTGTTACGACCATCCGTCATCTCTGGAGAGGTTTGGGAAGACCGGTGAGAA 994
QY 459 GCTGTACGGCCCGCTCGCGCGCTGACTACCTGGACAAACCAAGCGCTTCGCCCTGTT 518
DB 995 GATCTACGGACCTGACACTGGAGTTGATTACAAAGACCAACAGATCGGTTTCAGCTTCT 1054
QY 519 CTGCAAGCGCGTATTGAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564
DB 1055 TTGCCAGGACGACTCGAGGCTCTAGGATCTTAACCTTCAACAAACCAACCTACTTCAA 1114
QY 565 -----GGCAGGACTGGCTCTTCTGTCGCAAGACTGSCACTCGCCCTGCTGGCC 614
DB 1115 AGGAACCTTATGTGAGATGTTGTGTTCTGTCGCAACGACTGGCAGCTGGCCACTGGC 1174
QY 615 CGTCTGCTGAGGAGGAGTACAGCCCAAGGGCCAGTTTCAACCAAGGCCAAGTCGGTGT 674
DB 1175 GAGCTACCTGAAGAACAACTACCAAGCCCAATGSCATCTACAGGAATGCAAGGTTCTTT 1234
QY 675 GCGTATCCACATCGCTTCCAGGGCGCGATGTTGGGAGGAGGCTTTCAGGACACAGAA 734
DB 1235 CTGATCCCAACATCTCTTACCAAGGCGGTTTTCGCTTTTCGAGGATTTACCTGAGTGAA 1294
QY 735 GCTGCCCCAGGCGCTTTTGACAAGCTGCGCTTCTCGGACGCTATGCCAAGGTTTACAC 794

DB 1295 CCTCTCCGAGAGGTTTCAGGTCATCTCTGATTTCATGACGGGTATGAC----- 1343
QY 795 TGAGGCCACCCCATGGAGGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 854
DB 1344 -----ACGCGGTGGAGGCGAGGAA 1363
QY 855 GATCAACTGGCTGAAGGTGGCAATTATCGCCGCGGACAAAGCTGGTGTGCTGCCCAA 914
DB 1364 GATCAACTGGATGAAGCGCGGAATCTTGGAGCCGACAGGAGGCTCACCCTGAGCCGTA 1423
QY 915 CTAGCGACCGAGATCGTCCGATCGCCGCGGTGTGGAGTGGACACCGTCAATCCG 974
DB 1424 CTAGCGGAGGAGCTCATCTCCGATCGCCAGGAGTGGAGTGGACACATCATCGC 1483
QY 975 CGCAAGGGCATTCAGGGCATTCGAAGCGATTCGAGATTCAGGAGTGGACACCAAGAC 1034
DB 1484 GTCACCGGATCATCCGCGATTCGAAGCGATTCGAGTTCAGGAGTGGATTCCTAGCAA 1543
QY 1035 CGACAAGTTCCTGTCTGCGCCTACGACCCAGAACAGCGCTCTACCGCGGCAAGCGCGCC 1094
DB 1544 GGACAAGTATCATCACCGCAAGTACGACGCAACACCGCAATTCAGGCGAGGCGCTGAA 1603
QY 1095 CAAGGAGGCGCTGAGCGCGAGCTGGCGCTGCTGTGGACCCCGCCCTGTTTGGC 1154
DB 1604 CAAGGAGGCGCTGAGCGCGAGGCGGCTTCCGCTGCGAGGAGGAGGAGGAGGAGGAG 1663
QY 1155 CTTTCATCGCGCGCTGAGGAGGAGGAGGAGTGTGGACATCATCTGCGCGCGCTGCCCAA 1214
DB 1664 GTTCACTGCGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1723
QY 1215 GATCTCTGGCCACCCCGCAAGTTCAGATTCGCTGGTACCGGCAAGCGCGCTACGA 1274
DB 1724 GCTCATG---CAGGAGGAGCTCCAGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1780
QY 1275 GAGCTGGTGAACCGCATTCGCGCAAGTACAGGAGGCGCGCCGCAAGGCGGTGGTCAAG 1334
DB 1781 GAAGCTGCTCAAGAGCATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1840
QY 1335 CTGCGGCGCTGCGGACATGCTACCGCGCGCGGCGGCTTCTATGCTGCTGCTGCTGCTG 1394
DB 1841 CAAGCGCGCTTGTCTCATCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1900
QY 1395 CTTTCAGCGCTGCGGCTGATTCAGCTGCGACGCGCATGCTACCTACCTGCTGCTGCTGCT 1454
DB 1901 CTTTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1960
QY 1455 AGCTTCCAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1514
DB 1961 CGGTCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2020
QY 1515 CGGCTTGAACCCCGAC-----AAGCTGGAGGAGGCTGAGCGCGGCGGCGCGCGCCAC 1568
DB 2021 CGCTCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2080
QY 1569 GTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1628
DB 2081 CTTGAAGCGCGCATCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2140
QY 1629 CATCAGCGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1688
DB 2141 CATGAACCGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2200
QY 1689 GGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1720
DB 2201 CTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2232

RESULT 3
US-09-388-743-21
; Sequence 21, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:

```
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Ian
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; TITLE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 2067
; TYPE: DNA
; ORGANISM: Tulipa fosteriana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58)...(1857)
US-09-388-743-21

Query Match      13.2%; Score 281.2; DB 4; Length 2067;
Best Local Similarity 51.9%; Pred. No. 3.1e-39;
Matches 834; Conservative 0; Mismatches 678; Indels 96; Gaps 5;

QY 143 CTGTTACTGGTGCACCTGGCGCTGGACATCGTGTGCTGCTGAGG 202
DB 242 CTGTGCTTGGCTACGGTTCGAACACTACAGGATGAACCTGGTTCGTCGGGACCGAGA 301
QY 203 TCGCCCTTGGTCCAAAGACGGCGCTGGCGGATGTGACTGTGGTGGCTGCCCTATTGAGC 262
DB 302 CGGTCCTGTACAGACAGACGGCGGCTCGGGATGTGTAGGAGGTACCGCCGCC 361
QY 263 TGGTCAAGCGCGGCCACCGCGTCATGACCATTTGCCCTCGCTACGACCTAGTACGCTGAGC 322
DB 362 TGGCGCGAGAGGCGCATCGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
QY 323 CTGTGGACACCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 370
DB 422 CATGGGACACAACTGCTGTGATCAAGTCGGGATCAAGTCGGGATCAAGTCGGGATCAAGTCGG 481
QY 371 ACTTCCATCCATCAAGAGGCGTGCACCGCTGTGATGATGATGATGATGATGATGATGATGATG 430
DB 482 TCTTCCATCCATCAAGAGGCGTGCACCGCTGTGATGATGATGATGATGATGATGATGATGATG 541
QY 431 CCAAGCTTGGGCGACAGCGCTGCAAGCTGTACGGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 490
DB 542 AGAAGCTTGGGCGAAACCGTGGGAAATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
QY 491 TGGACAAACCAAGCGCTTGGCGCTGCTGCAAGCGCGCTGATGAGGCTGCGCGGCTGCG 550
DB 602 ATGATACCAAGTACCGTTCAGTCTTCTGTGTCAGGCTGCTTGGAGGCTCCAGAGTTC 661
QY 551 TGGCTTGGCGGCC-----GGCAGGAGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 586
DB 662 TAAATCTCAACACAGTAAATATTTTCTGGACCATATGTTGAAGATGCTGTTTTATG 721
QY 587 CCAACAGTGGGACTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 646
DB 722 CGAACGATTTGGCACTGACCTCTTCCATGCTTACTTGAAGAGTGTATATAATCAAG 781
QY 647 GCCAGTTTCCAAAGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 706
DB 782 GATTATATGAGAGTGCAGAGTGGCTTTTGCATTTCAATATATGCAATACCAAGGAGAT 841
QY 707 TGTGGAGGAGGCTTTCAAGACACGAGCTGCCCGGCGGCTTTGCAAGCTGGCTTTGCAAGCTGGCT 766
DB 842 TCGCTTTCTGATTTCTGCTTCTCAACCTTCCAGACATTTAAATCTGCTGTTGATT 901
QY 767 TCTGAGGCTATGCAAGCTTTTACACTGAGGCCACCCCTGAGGAGGAGGAGGAGGAGGAGGAGG 826
DB 902 TCTTCGATGATATACAAAACCT----- 924
QY 827 CCGGCTGAGGGAAGACCTACAAGAGATCAACTGGCTGAAGGTTGCAATTATCGCG 886
DB 925 -----GTGAAGGTAGAAAAATAATTGGATGAAGGCTGGATATTTGAAG 970
```

```
QY 887 CCGACAAGCTGGTACTGTGTGCCCAACTACCGACCGAGATCGTGTGCCGATGCGCGCG 946
DB 971 CGGACACTGTGTAACTGTGAGCCCGTATTATGCTTAAGAGCTGCTCTCTGGAGAAGATA 1030
QY 947 GCGGTGTGAGCTGGACACCGCTATCCGCGCAAGGCGCATTTAGGCGCATTTGCAAGCGCA 1006
DB 1031 GAGGTGTGAGTGGACAACGTTCTCGGCTTGGAGGCGCTCAAGAGAAATTTGTAATGGGA 1090
QY 1007 TGGACATTTGAGGAGTGAACCCCAAGACCGACAAAGTTCCTGTCTGCGCCCTACGACGACA 1066
DB 1091 TGGATACTAATGTGTGATTCATTCACAGACAAATTTATCACTGCAAAATTCAGATCAA 1150
QY 1067 ACAGCGCTTACGCCGCAAGCGCGCCCAAGAGGCGCTGTCAGGCGGAGCTGGCGCTGC 1126
DB 1151 CAATGTAACAGAGGCAAAACGTTTAAATAAGCAAGAAATTTACAAGCAGAGTTGGCTTGC 1210
QY 1127 CTGTGACCCCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1186
DB 1211 CTGTGATCCAGACATTCCTGTTATAGTTTTTTTGAAGGCTGGAGGACGAGAAAGTTT 1270
QY 1187 TGGACATCATCTCTGCGCGCTGCTGCCCAAGATCTGCTGCCACCCCAAGGTGACATCGCCA 1246
DB 1271 CAGATATCTAGCTGCAGCAATTCAGAAATTAATGG---ATGAGAAGCTTCAGATCATAA 1327
QY 1247 TCGTGGTACCGGCAAGCGCGCTACAGAACTGTGTGACGCGCTGTCAGGCGGACCAAGTACA 1306
DB 1328 TTCTTGAACCTGCAAGAAACACCTCGAAAGGAGCTTGAAGAAATAGAAAGAACAAATTC 1387
QY 1307 AGGCGCGCCCAAGGCGTGTCAAGTTCCTGCGCGCTGCGCGCACATCTCACCGCG 1366
DB 1388 CAGACAGATGAGACTTTTCCGAAATTAATGTTCGTTGGCTCATATGATGATGCTG 1447
QY 1367 GCGCGACATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1426
DB 1448 GAGTGATTTTATAATTAATTCCTAGTAGATTTGAGCGCTGCTGGCTTATTTCAGCTTGAAG 1507
QY 1427 CCATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1486
DB 1508 GCATGAATATGAGGATGCCAGCCATATGTTCCACACCGGCTGCTGCTGCTGCTGCTGCTG 1567
QY 1487 AGGAGGCGTCAACGCTTCCATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1540
DB 1568 GGAAGGCTTACCGGATTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1627
QY 1541 AGGCTGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1600
DB 1628 CGGTGATGCTGCGGGAATTTTAAACTGTAAGAGGCGCTTAAAGTCTATGGAATC 1687
QY 1601 GCGCTACCGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1660
DB 1688 CAGCCTTACGGAATGTTTCAAGAACTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1747
QY 1661 CCAGAAAGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1708
DB 1748 CAAAAAATGGAGAGAGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1795
```

RESULT 4

US-09-388-743-17

```
; Sequence 17, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Ian
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; TITLE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
```

; LENGTH: 2274		
; TYPE: DNA		
; ORGANISM: Typha latifolia		
; FEATURE:		
; NAME/KEY: CDS		
; LOCATION: (115)...(1956)		
US-09-388-743-17		
Query Match		
Best local similarity 12.6%; Score 266.6; DB 4; Length 2274;		
Best local similarity 52.9%; Pred. No. 8.6e-37;		
Matches 828; Conservative 0; Mismatches 639; Indels 98; Gaps		
QY	172	GCCTGTGCACATCGTGTGATGTTGCTGTGTGAGGTGCGCCCTTGTTGTCGAAGACGGCGGCGCTG 231
DB	373	GGGATGAACCTAGTCTTTGTGTAGCTGAGATGGCTTCCATGGAGCAAGACTGGAGGCTT 432
QY	232	GGCGATGTGACTGGTGGCCCTATTGAGCTGTGTCAACGGCGGCCACCGGTCATGACC 291
DB	433	GGTGTATGTTCTTGGAGGACTCCACCGGCAATTGGCGGCAAAATGGACATCGAGTTATGTT 492
QY	292	ATPTGCCCTCGCTACGACCAAGTACGCTGACCGCTGGGACACCTCGGTGGTTCGTGACAT- 350
DB	493	ATAGCGCCACGTTATGATCAATACATGATGCTTGGATACAGATGCCTCTGTTGAGTTG 552
QY	351	-----CATGGCGAGAGGTCGCGTACTTCCATCCATCAAGAAGGGCGTGCAC 399
DB	553	AAAGTTGGGATAGGTTGAAACCGTCGCGTCTTTCTACTGTATATAAAGAGGAGTTGAT 612
QY	400	CGCGTGTGGATTGACCACCCCTGGTTCCTGGCCAGGCTCTGGGCAAGACCGGCTCCAAG 459
DB	613	CGAGTTTGTGCAATCACCTATGTTTCTTGCAGAGCTGCGGGGAAACTGGTGGGAAG 672
QY	460	CTGTACGGCCCCGCTCCGGCGGTACTACCTGGACAACCAAGCGGTTGCGCCCTGTTTC 519
DB	673	ATTATATGGTCTTAACACTGGAACAGACTATCAGGACAATCAGTACGCTTCAGCTTTC 732
QY	520	TGCAAGGCGCTATTGAGCGTCCCGCGTCTGCCCTTCGCCCTTCGGGCGCC----- 564
DB	733	TGCCAGCAGCATTTGGAAGCTCTAGAAATCTTAAATCTCAACAACAGTGAITCTTTCTCT 792
QY	565	-----GGCGAGGACTGGGTCTTCGTGGCCACGACTGCACCTCCGCCCTGGTGCC 615
DB	793	GGTCCATTATGGGAAGATGTAATCTTCATTTGCATATGATTGGCAGACTTCCTCTCGCA 852
QY	616	GTCTGTCTGAAGCAGGATACCAAGCCCAAGGCCCAGTTTCAACAAGGCCAAAGTCGGTG 675
DB	853	TGCTACTTAAAGAGCATGTACCATCCCGTGGCAATTTACAAGAAGCCCAAGGTTGCTTC 912
QY	676	GCTATCCACACATCCGCTTCCAGGCCCCCATGTGGGAGGAGGCTTTCAGAGCAACAGAG 735
DB	913	TGCATTACAAATATATCATACCAAGTCGATTTTCCCTCAGACTTCGAATTTCTCAAT 972
QY	736	CTGCCCCAGGCGCCCTTTGACAAGCTGGCTTCTCGGACGGCTATGCCAAGGTTTACACT 795
DB	973	CTTCCGGAG-----AATTCAATCT 993
QY	796	GAGGCCACCCCATGAGGAGGACGAGAGCCCGGCTGAGGGAAAGACCTTACAAGAAG 855
DB	994	TCCTTCAGTTTTCATTATGGGTACAACAAGC-----TGTAAGGGATGAAG 1041
QY	856	ATCAACTGGCTGAAGGGTGGCAATTATCGCCCGGACAGCTGGTGACTGTGTGCGCCCAAC 915
DB	1042	ATAAATTGGATGAAGCAGGAATCTTAGAATCAGACGGGTGTTTACAGTAGTGCATAT 1101
QY	916	TACGCCACCGAGATCGTCCGATCGCGCGGTGTGGAGCTGGACACCGTCAATCCGC 975
DB	1102	TATGCACAGAGCTCTTTTCAGSAGAAGAAAGGGAGTCGAGTTGGACAACATTTTCGGT 1161
QY	976	GCCAAGGCCATTGAGGGCATTTGAACGSCATGGACATTGAGGAGTGGAAACCCCAAGACC 1035
DB	1162	GTGACCAGTATCACAGGAATTGTGAATGGAATGGAATTTAATGAGTGAATCCATTACA 1221
QY	1036	GACAAGTTCTGTCTCGGCCCTACGACCCAGAACAGCGTCTACGCCGCAAGGCGCGCC 1095

Matches 795; Conservative 0; Mismatches 649; Indels 96; Gaps 5;

```
QY 176 TGGACATCGTGTGTTGCTGAGTGGCGCCCTTGGTCCAAAGACGGCGGCGTGGCG 235
Db 349 TGAACITGGTGTGTTGTTGAGTGTAGTCCATGGAGCAAACTGGGGCGCTTGGCG 408
QY 236 ATGTGACTGGTGGCCCTATTGAGCTGTTCAAGCGCGGCCACCGCTCATGACCATG 295
Db 409 ATGTTCTTAGAGGATTGCCACCTGCTATGCTGCAATTTGGGCACAGGGTCATGACCGTGG 468
QY 296 CCGCTCGCTACGACCAAGTACGCTGAGCGCTGGGACACCTCGGTGGTTCGTGGACAT----- 350
Db 469 TGCCAGATATACCAATAAAGATATCTGGGATAAAGTGTCCAGTGTAGTTAAAG 528
QY 351 -----CATGGCGAAGGTCGCTACTTCCACHTCCATCAAGAAAGGGGCGTCCACCGG 403
Db 529 TTGGGGATAGATTGAACCTGTCGGTCTTCCACTGCTACAAAAGGGGATGATCGGG 598
QY 404 TGTGATTGACCAACCCCTGTTCTGCGCAAGGTCTGGGCAAGACCGGCTCCAAAGCTGT 463
Db 589 TTTTGTGGATCACCTATGTTTTCGAGAAGGTTTGGGGGAAAACAGAGGAAATAT 648
QY 464 ACGGCCCCCGCTCGGCGTGTACTTGGACAAACACAGCGCTTGGCCCTGTTGCA 523
Db 649 ATGCTCTGTACAGAAACAGATATTCAGACAAATCAACTAAGATTACGCTTTTGTGC 708
QY 524 AGCGCGCTATTGAGGCTGCCGCGTGTGCTGCTCGCCCGCGGAGGACTGC----- 576
Db 709 TGCAGCTGTGAAGTCTCAAGACTTCTAATCTCAACACACAACTATCTTCTGGAC 768
QY 577 -----GTCTTCGTGGCCAAAGCTGGCACTCCGCCCTGGTGGCGCGTCC 619
Db 769 CATATGGAGATGATGTTGTGTTTATGCGCAAGATTGGCATTTCTGCTACTGCGCGCT 828
QY 620 TGCCTGAGGAGTACCAACCCCAAGGCGAGTTCACCAAGGCCAAGTCTGGTGTGCTA 679
Db 829 ACTTGAACATATGTACCAATCATGTTATTTACATGAATGTAAAGTTGSCATTTTGA 888
QY 680 TCCACAACTCGCTTCCAGGCGCGATGTGGAGGAGGCTTTCAGGACACGAAAGTGC 739
Db 889 TTTATATATTGCTTACAGGCGCGATTTGCTTTTCGGACTTGAACCTCTTAATCTCC 948
QY 740 CCGAGCGCGCTTTGACAGCTGGCTTCTCGGACGCTATGCCAAGTTTACACTAGG 799
Db 949 CCAATAAATTTAAATCTTCAATTTGATTTCAATGATGATATGACAAACCT----- 998
QY 800 CCACCCCATGGAGGAGGACGAGAGCCCGCTGACGGGAAAGACCTCAAGAAGATCA 859
Db 999 -----GTGAAAGGAGGAAATAA 1017
QY 860 ACTGGCTGAAGGTTGGCATTTATCGCCGCCGACAAAGCTGTGCTGTCGCCCACTACG 919
Db 1018 ATTGGATGAGGCTGAATAATAGATGTATAGTGTGCTTACCGTGTAGGCCCATATTATG 1077
QY 920 CGACCGAGATCGTGTGCGGATGCGCGCGGTGTGGAGCTGGACACCGCTCATCGGCCCA 979
Db 1078 CCAAGAGCTGTCTCAGGCTGAGAAAGGTTGTGAGTTGGGCAATATCTGTGCGCATGA 1137
QY 980 AGGSCATTGAGGCGATTGTGAACGGCATGGACATTTGAGGAGTGTGAACCCCAAGACCGACA 1039
Db 1138 AAACCATCTGTGGAAATAGTAATGGGATGACACACCGAGTGGAAATCAATTAACAGACA 1197
QY 1040 AGTTCCTGTCTGCGCCCTACGACAGAACACGCTTCTACCGCGGCAAGCGCGCGCCCAAG 1099
Db 1198 AATATATTCTACAAACTACGATCAACAACCTGATTGATGATGATGATGATGATGATGATG 1257
QY 1100 AGCCCTGTGAGGCGAGCTGGGCTGCTGTGACCCCGCCCGCCCTGCTGCTGCTGCTGCTG 1159
Db 1258 AAGCTTTTGCAGCTGTGAGTGTGGGCTGCTGTTAAACAAAACAAAGCTTGTGTTGGCCCTTG 1317
QY 1160 TCGGCCCTGTGGAGGACGAAAGGTTGTGGACATCATCTTGGCGCGCGCTGCCCAAGATCC 1219
Db 1318 TTGGAAGACTAGATGACGAAAGGCTCAGACATTTCTAGCTGACGAAATTCAGAACTTC 1377
```

```
QY 1220 TGCCACCCCAAGTGTGAGTCCCATCTGTTACCGCAAGCGCGCTACGAGAAGC 1279
Db 1378 T---TTTGAGAAATGTTCAAGTGTATGACTTGGCACTGCGCAAGAAAGTTGGAGTG 1434
QY 1280 TGTGAAGCGCATCGGACCAAGTACAAGGCGCGCCCAAGGGGTGGTCAACTTCTCGG 1339
Db 1435 AACTTCAATTACTTTGAGGAAATGTTTCCAGACAAATTCAGAGCACATCTCAATTCACG 1494
QY 1340 CGCCCTGGCGGCATGCTCACCGCGCGCGGCTCATCTGCTGGTGGCTCGCGTTCG 1399
Db 1495 TTCTTTAGTCTATGCAATCATGCGAGGAGCTGATATCTCTTGTATTCAGACAGATTCTG 1554
QY 1400 AGCCCTGCGGCGCTATCCAGCTGCACGCCATGACATACGTTACCGTGGCGCGTGGTACCT 1459
Db 1555 AACCTGTGCGCTCATTCAGCTTCAGGCCATGCGATATGAACTCTCCCTATGTGTAGCA 1614
QY 1460 CACCGCGCGCTTGGTGCACACCGTCAAGGAGGCGCTCACCGGCTTCCACATGGCGCC- 1518
Db 1615 CCACTGTGAGCTTGTTCACACTGTCAAGAAGGCTTCACTGGCTTCCATATGGGCCCCCT 1674
QY 1519 -----CTGAACCCCGCAAGCTGGACGAGCTGACGCGGACGCGCTGGCGGCCACCGTGC 1573
Db 1675 TCAGTGTGGAGTGTGCTGCGGTAGACAAAGCTGATGTACAAAGATTGTGAAACACGGA 1734
QY 1574 GCGTGTCCAGCGAGGTGTTTGGCGGCGCGCTACCCCGAGATGTTGGCCAACTGTCATCA 1633
Db 1735 AAAGGCGCCCTCAAAGTCTATGGAACACCTGCTTTGTGGAGATGATCAAGAAGTGCATGA 1794
QY 1634 GCAGAGCTGTCTGTTGTCACAGCCCGCCAGAGTGGGA 1673
Db 1795 ACCAAGATCTCTATGGAAGGACCTGCAAGAAGTGGGA 1834

RESULT 6
US-09-388-743-1
; Sequence 1, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; FILE OF INVENTION: 1144
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2202
; TYPE: DNA
; ORGANISM: Curcuma zedoaria
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (130)...(1974)
US-09-388-743-1

Query Match 12.0%; Score 255.4; DB 4; Length 2202;
Best local similarity 51.8%; Pred. No. 6.5e-35;
Matches 799; Conservative 0; Mismatches 646; Indels 96; Gaps 6;
```

356 -----GCGAGAGCTCCGCTACTTCCACTCCATCAAGAGGCGTGCACCGG 403
Db 575 TTGGTGATAGATTGAAACTGTTGCTTTTCCACTGCTACAAAGAGGGAGTTGATCGGG 634
QY 404 TGTGATGACACCCCTGTTCTGGCCAAAGTGTGCGGCAAGACCGGCTCCAAAGCTGT 463
Db 635 TGTGTTGGATCACCTCTCTCTCTGAGAGGTTTGGGAAAAAATCGAGGAAGATAT 694
QY 464 ACGGCCCGCTCGGGCGCTGACTACCTGGAACACCAAGCGGTTCGCCCTGTCTGCA 523
Db 695 ATGGTCTGTGCACAAGAACTGATTGAAGACAAACAGTAAAGTTCTGTCTGTCTGT 754
QY 524 AGGCGCTATTGAGGCTGCGCGCGTGTG-----CCCTTCGGCC 562
Db 755 TGGCAACTGCGAACTCCAGGGGTTCTGAATCCCAACAATAAATATCATTTCTGGAC 814
QY 563 CCGCGGAGGACTGCTGTGCGCCAAAGCTGCGACTGCGCTGCGCTGCGCTGCTGCG 622
Db 815 CAAAGGTGAAGATTTATTCATGCTAACGATTGGCATGCTCTATTACCTTGCTATT 874
QY 623 TGAAGGAC---GAGTACAGCCAGGCGCAGTTTCCACCAAGGCGCAAGTGGTGTGCTA 679
Db 875 TAAAGACCATGTATATCAAGCCCATGGAATACAAATAATGCTTAAAGTTGCTTCTGCA 934
QY 680 TCCACAAATCGCCTTCCAGGCGCGCATGTGGGAGGAGGCTTTCAGAGGACACGAAGTGC 739
Db 935 TCCATAATATTGCTATCAGGACGCTTGGCTTTGAAGATTTTCGGTCTCAATCTCC 994
QY 740 CCAGGCGGCTTTGACAAGCTGCGCTTTCGGAAGCTATGCGCAAGTTTACACTGAGG 799
Db 995 CTGATACATTCAAGTCTCTTTGATTTCATGATGCTATGCAAA-----1040
QY 800 CCACCCCATGAGGAGGACGAGAGCCCGCTGACGGGAAGACCTACAGAAGATCA 859
Db 1041 -----ACCAATAAAGGAAGAAATCA 1063
QY 860 ACTGGGTGAAGGTGGCATTTATCGCGCGCAAGCTGCTGTCGCCCAACTACG 919
Db 1064 ACTGGATGAAGGGGAATTATAGATCAGATGCTGCTGCTGAGGCCATACHTATG 1123
QY 920 CGACCGAGATCGCTGCGGATCGCGCGGCTGTGAGCTGGACACGCTCATCGCGCCA 979
Db 1124 CCAGGAACCTGCTCTCAGGAATCGATAAGGCGCTGAGTTGGACAATATCTGCGCTGA 1183
QY 980 AGGCGATTGAGGCACTGTGAAGCGGATGACATTTGAGGACTGGAACCCCAAGCAGCA 1039
Db 1184 AAACCATCTGTGCATCATAAATGGAATGGAACACACAGAGTGGATCCCTCAACAGCA 1243
QY 1040 AGTCTGCTGCGCCTACGACCAAGACAGCTCTACGCGGCAAGCGCGCGCAAGG 1099
Db 1244 AATACATAACAGAAATACGAGGCAACCACTGTAATGGAGGCAAGGCCACTCAACAGG 1303
QY 1100 AGGCGCTGAGGCGGAGCTGGGCTGCTGTCGACATTCCTGCGCGCTGCCCAAGATCC 1159
Db 1304 AAGCTTGGCAAGCTGAGGTTGAGCTGCGCGTCAACAGTAAATCCCTGTGATAGCTTCA 1363
QY 1160 TCGGCGCTGAGGAGCAGAGGCTGTGACATTCCTGCGCGCTGCCCAAGATCC 1219
Db 1364 TTGGCAGACTAGAACCAAAAGGGTTGACATTTCTAGTGAAGCAATTCCTCAAGTTCT 1423
QY 1220 TGGCCACCCCAAGGTGAGATGCGCATCTGGGTACCGGCAAGGCGCGCTACGAGAAGC 1279
Db 1424 T---CGATCAGGATGTTCAAGTGATAGTTCTCGTACTGTTAAAGAAAGATTACGCGC 1480
QY 1280 TGGTGAACGCCATCGCACCAAGTACAGGCGCGCGCCAGGCGGTGGTCAAGTTCTCGG 1339
Db 1481 AACTTGCATGCTCGAGGACGAGTTCCCAAGACAAATTCAGAGCTCATATGAAGTTCAATA 1540
QY 1340 CGCCCTGCGGCACATGCTCACCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1399
Db 1541 TTCTCTTGGCTCATGAATCATGCGGCGGCTGATATCTCTGTTATTCCCAAGTAGTTCG 1600
QY 1400 AGCCCTGCGGCTGATCCAGCTGCAGGCGATGCTACGCTACGCTGCGGCTGCTGCTGCT 1459

Db 1601 AACCATGGGCTCTCATTTAGCTTCCAGGCGATGAGATACGGAACCCCTTCCATGTGCACCA 1660
QY 1460 CCACCGGCGGCTGCTGACACCGCTCAAGGAGGCGGTTCACCGGCTTCCACATGGGCGGC- 1518
Db 1661 GCACTGGTGGGCTCTGCGACACTGTCAAAGAGGCGATCACAGTTTCCACATGGTCCCT 1720
QY 1519 -----CTGAACCCCGCAAGCTGGACAGGCTGACGCGCGACGCGCTGCGCGCACCGTGC 1573
Db 1721 TCAGTGTGAGTGGACATTTGCCGACGAGCGCGAGCTGCTAAAGATTGGAAGCAGTGA 1780
QY 1574 GCCGTGCCAGCAGGTGTTGGGGGCGCGCTACCCCGAGATGTTGCCCAACTGCATCA 1633
Db 1781 AGAGAGCCCTTATGCTTTATGGAAGCGCTGCTTTCGAGGAGATGATACAGAACTGCATGG 1840
QY 1634 GCCAGGACCTGCTCTGGTCCCAAGCGCGCCAGAAAGTGGAG 1674
Db 1841 CTCAGATTTCTCTGGAAGGCGCGCAAGAAAGTAATGGAG 1881

RESULT 7

US-08-941-445A-10
; Sequence 10, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2097 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2097
; US-08-941-445A-10

Query Match 10.3%; Score 219.2; DB 3; Length 2097;
Best Local Similarity 50.6%; Pred. No. 7.4e-29;
Matches 801; Conservative 0; Mismatches 673; Indels 108; Gaps 7;

144 TGTACTGTGCGCACTGGTGGCCACTTGGCGCTGGACATCGATGATGGTCTGCTGAGGT 203
Db 588 TGAACCTGGCCCTTTGGCTGGGCGCTAAATGTAAGACGTCGTCGGTGGCTTCTGAATG 547
QY 204 CGCCCTTGGTCCAAAGACGGGCGGCTGGCGGATGTGACTGTGGGCTTGCCTATTGAGCT 263
Db 648 TGTCTCTTCTTCAAGACAGGTGGCTTGGAGATGCTGGTGGTCTTTCCTTAAGGCTCT 707
QY 264 GGTCAAGCGCGCCACCGCGTCAATGACCAITGCCCCCTCGCTAGCAGCAAGTACGCTGACGC 323
Db 708 GCGAGGAGAGACACCGTGTATGCTGCTGATACCAAGATATGAGAGATATGCCGAAGC 767
QY 324 CTGGGACACTCGGTGGTGGTGGACATCATGGCGAGAGAGTCCGCTACTTCCACTCCAT 383
Db 768 CCGGG--ATTAGGTAAAGGAGCTTACAAAGGTATGGAGAGATTCAGAAATTA 825
QY 384 CAAGAAGGCGTGCACCGGCTGTGGATTGACCAACCGCTGTCTGCGCAAGGCTGGGG 443
Db 826 TATTTTCACTTTACATTTGATGGATTGATTTTGTATTCGTAGAACCGCTTCCCTTCGG 885
QY 444 CAAAGCCGCTCCAGCTGACGGCCCCCGCTCCGCGCTGACTACTGGACAAACCAAA 503
Db 886 CAC-----CGGCACAAATAATTTATGGGGGAGAAAGATTGGATATTTTGA 932
QY 504 GCGCTTCGCGCTGCTGCAAGCGCGCTATTGAGGCTGCCGCGTGTGCTGCGCTTCGCGCC 563
Db 933 GCGCATGATTTTGTCTGCAAGCGCGCTGTTGAGGTTCCATGATGCTCCATGTGGCG 992
QY 564 -----CGGCGAGACTGCTCTTCTGTTGGCAAGCTGGCACPCGCCCTGGT 611
Db 993 TACTGTCTATGTTGATGGCACTTAGTTTTCATTTGATGATTTGATGATGCTGCTGCTGT 1052
QY 612 GCGCTCTCTGCTCAAGGACGATACCGCCCAAGGCGCTTCCACCAAGGCGCAAGTCGT 671
Db 1053 GCGTGTCTATCTAAGGCGCTTACCGGGACAAATGTTTATGCGAGTATGCTGCTGCTGT 1112
QY 672 GCTGGCTATCCAAACATGCGCTTCCAGGCGCGCATGTGGAGGAGGCTTTCAAGGAC 731
Db 1113 GCTGTGATACAAACATTTCTCATCAGGCTGTCGCGCTGTAGACGACTTCTGCAATTT 1172
QY 732 GAAGTGTCCCGACCGCGCTTTGACAAAGTGGCGCTTCTCGAGCGGCTATGCCAAGTTTA 791
Db 1173 TGACTTGCCTGAACACTATACATCGACCTTCAAACTGTATGACAACTATGTGGGGATCA 1232
QY 792 CACTGAGGCCACCCCATGAGGAGGAGAGAGACCGCCCTGACGGGAAAGACCTACAA 851
Db 1233 C----- 1233
QY 852 GAAGATCAACTGCTGAAGGTTGGCATTTATCGCGCGGACAAAGTGTGTGTCGCC 911
Db 1234 ---AGCAAGTWTTCGTCGCGGCTGAAGACGGCAGACCGGCTGTGACCGTTAGCAA 1289
QY 912 CAACTAGCGACGAGATCGCTCGGATGCCGCGCGGCTGTGGAGCTGGACACCGTCAT 971
Db 1290 TGGCTACATGTGGAGCTGAAGACTTTCGGAAGCGGCTGGGCGCTCCACGACATCAAA 1349
QY 972 CC---GGCCCAAGGCAATGAGGCGCTTGAACGGCATGGACATTCAGGAGTGGAAACC 1028
Db 1350 CCAGACGACTGGAAGTGCAGGCGCATCTGTAACGGCATTCGACATGAGCGAGTGGAAACC 1409
QY 1029 CAAGACCGCAAGTTCCTGTCTGCGCCCTACGACCAAGACAGC-----GTCTA 1076
Db 1410 CGCTGTGACGCTGACCTCCACTCCGACGACTACCACTACCACTACAGTTCGAGAGCGCTGA 1469
QY 1077 GCGCGGAAGCGCGCGCGGAGGCGCTTCAGGCGGCTGCGCTGCTGTGGACCC 1136
Db 1470 CACCGGAAGCGGAGTGCAGGCGCGCTTCAGCGGCGCTGCGGCTGCGAGTCCCGGA 1529
QY 1137 CACGCGCCCTGTTGCTGCTTATCGGCGCTGAGGAGGAGGAGGAGTGTGGACATCAT 1196
Db 1530 CGAGTGCCTATGCGGTTCTCGGCGCTGACCGCGCTGGACCAAGAGGCGTGGACATCAT 1589

QY 1197 CTTGGCGCGCTGCCCCAAGATCTGGCCACCCCAAGTGTGCGATCGCCATCTCTGGGTAC 1256
Db 1590 CGCGGACGCGATCACTTGGATC---CGGGGAGGAGCTGCTGATGCTGGGCAC 1646
QY 1257 CGCAAGGCGCGCTAGCAGAACTGTGACGCACTGGCACAAGTACAAAGGCGCGCGC 1316
Db 1647 CGGGCGGGCGGACCTGGAGGACATGCTGGCGGCTTCGAGTGGAGACAGGCAAGGT 1706
QY 1317 CAAGGCGGTGGTCAAGTCTCGGCGCGCTTGGCGCACATCTCACCGCGCGCGCGCACTT 1376
Db 1707 GCGCGGTGGTGGGTTCTCGGTGCCCTGGCGCACCGCATCACGGCGGCGCGGACAT 1766
QY 1377 CATCTGTGTGCTCGCTCGCTTTCAGCCCTGCGGCGCTGATCCAGCTGCACGCCATGACTA 1436
Db 1767 CTTGTGTATGCTGCTGCGGTTTCGAGCGCTGCGGCTGAACCACTCTACGCCATGCGTA 1826
QY 1437 CGTACCGTGGCGCTGTAGCTTCCACCGCGCGCTGGTGCACACCGTCAAGGAGGCGCT 1496
Db 1827 CGGACCGTGGCGCTGGTGCACCGCTGGGGGGCTCCGGGACACCGTGGCGCGCTTCCA 1886
QY 1497 CACCGGCTTCCATGGCGCGCTGAACCCCGACAACTGGAGAGCTGAGAGCTGACCCGACGC 1556
Db 1887 CCGGTTCAAGCACACCGGCTCGGTGGAGCTTTCGACCGCGCGGAGCGCAACCGGATGAT 1946
QY 1557 CTTGGCGCGCACCGTGGCGCGCTGCGAGGAGTGTGCGGGGCGCGCTACCCCGAGAT 1616
Db 1947 CGACGCGCTCTCGCACTGCTTCCACCGTACCGGAACTACAAGGAGAGCTGGCGCGCTG 2006
QY 1617 GGTGGCACTGATCAGCAGGACCTGCTGTCGTCGCAAGCGCGCGCGAGAGTGGAGGG 1676
Db 2007 CAGGGCGCGGATGCGCGAGGACCTCAGCTGGGACCAAGCGCGCTGCTGTATGAGGA 2066
QY 1677 CTTGTGGAGAGTGGTGTAC 1698
Db 2067 CTTGTCTCAAGCGGAATAC 2088

RESULT 8

US-09-196-390-5
; Sequence 5, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; APPLICANT: Lorz, Horst
; APPLICANT: Lutticke, Stephanie
; APPLICANT: Walter, Lennart
; APPLICANT: Froberg, Claus
; APPLICANT: Kossman, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
; TITLE OF INVENTION: SYNTHESIS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 21 588.9
; FILING DATE: 29-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 36 917.7

APPLICANT: KEELING, PETER L.
APPLICANT: KNIGHT, MARY E.
APPLICANT: GUAN, HANPING
TITLE OF INVENTION: MODIFICATION OF STARCH
TITLE OF INVENTION: SYNTHESIS IN PLANTS
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN
ADDRESSEE: Intellectual Property Group of
ADDRESSEE: Pillsbury Madison & Sutro LLP
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/572,951
FILING DATE: 15-DEC-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/346,602
FILING DATE: 29-NOV-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/263,921
FILING DATE: 21-JUN-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Paul N. Kokulis
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 222957/1.02.15C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2380 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-572-951-3

Query Match 10.0%; Score 213.4; DB 1; Length 2380;
Best Local Similarity 32.3%; Pred. No. 6.9e-28;
Matches 542; Conservative 239; Mismatches 810; Indels 86; Gaps 7;

QY 149 CTGGTCCCACTGTGCTGCGGCTGGACATCGTGTGCTGCTGAGTGGCC 208
DB 650 CNGNCCNTNGCNGGNCNNAAYGTWATGAAYGTNGTNGTNGCNGWNGARTGTCNC 709
QY 209 CTTGTGTCAGAGCGGCGCTGGCGATGTGACTGTGCTGCGCTGCTATGAGTGTCA 268
DB 710 CNTTGTGAARACNGGNGNYTNGGNGAYGTNGTNGGNGCNYTNCNNAARGCNTNGCNM 769
QY 269 AGCGGCGCACCGCTCATGACCATGTCGCCCTCGCTACGACACAGTACGCTGACGCTGG 328
DB 770 GNNGNGNCAYMNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 829
QY 329 ACACCTCGGTCGTGGACATCATGGCGGAGAGGTGCGCTACTTCCACTCCATCAAGA 388
DB 830 --AYTNGGNGTNGMNGMNGNTAYARTNGTNGCNGNCAR-GAYWNGARTNACNTAYT 886
QY 389 AGGCGTGCACCGCTGTGATGATGACACCCCTGGTTCCTGGCCCAAGGTGTGGGCAAGA 448
DB 887 TYCAYWSNTAYTHGAYGNGTNGAYTGTNTTNGTNGARGCNCNCCNTTYMGCAYM 946
QY 449 CCGGCTCCAAAGCTGACGGCCCCCGCTCCGGCGCTGACTACCTGGACACCAACGCT 508

DB 947 GNCAYAAAYATHTAYGNGNGARMGNTNGAYATHTYNAARMGNATGATHYNTT 1006
QY 509 TCGCCCTGTTCTGCAAGCGCGCTATTGAGCTGCGCGCTGCGCTTCTGGCGCCGCGG 568
DB 1007 GYAARGCNGCNGTNGARGTNCCTGTGTCNCNCTGTCGNGGNGCNGTNGTAYGNGAYG 1066
QY 569 AGGACTGCTCTTCGTGGCCACAGCTGGCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 628
DB 1067 GNAAYTNGTNTTYATHGCGNAAAYGAYTGCAACNCGNTNTTNCNGTNTAYTNAARG 1126
QY 629 ACAGTACAGCCCAAGGCGCTTCCACAGGCGAAGTGGTGGTGGTGGTGGTGGTGGTGGT 688
DB 1127 CNTAYTAYMNGAYAAAYGNYTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGT 1186
QY 689 TCGCTTCAGGCGCGCATGTGGGAGGAGCTTTCAGGACAGCAAGCTGCGCCAGCGCG 748
DB 1187 THGCNAYCARGNMGNGCNGTNGAYTGTNTYATNTYATNTYATNTYATNTYATNTYATNT 1246
QY 749 CTTTGACAAGCTGCTTCTCGGACGGCTATGCCAAGTGTACACTGAGGCGCACCCCA 808
DB 1247 AYATHGAYCAYTTYAARYTNTAYGAYATHTGNGGNGAYCAYWSNAAAYGTNTTYGC-- 1304
QY 809 TGGAGGAGGAGGAGAGCGCCCTGACGGGAAAGACCTACAAGAGATCAACTGGCTGA 868
DB 1305 ----- 1304
QY 869 AGGTGGCATATPCGCGCGGCAAGCTGTGTACTGTGCGCCCAACTACGCGACGAGAGA 928
DB 1305 -NCGNGTNTAARACNGCNGAYGNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 1363
QY 929 TCGCTCCGATGCGCGCGGCTGTGGAGCTGGACACCGCTATCCCGCGCAAGGCA--- 985
DB 1364 TNAARACNWSNGARGNGGTGGGNTNCAYGAYATHTAAYACAAAYGAYTGAARY 1423
QY 986 TTGAGGCGATGTGAACGCGATGAGATTCAGGAGTGGACCGGACCGACACGACGACGTC 1045
DB 1424 TNCARGNATHGTNAAYGNGTNGAYATGWSNGAR--GGAAYCNGCNGTNGAYGTNCAY 1483
QY 1046 TGTCTGCGCCTACGACCAAGC-----AGCGTCTACGCGCGCAAGCGCGCG 1093
DB 1484 TNCAYWSNGAYGAYTAYACNAAAYTAYACNTYARACNYTNGAYACNGNNAARMNCART 1543
QY 1094 CCAAGAGGCGCTGACGCGCGAGCTGGCGCTGCTGCGACCGCCCGCGCTGCTGCTGCTG 1153
DB 1544 GYAARGCNGCNYTNCARMGNCARYTNGGNTNCARGTNGMNGNGAYGAYTNCCTYATNG 1603
QY 1154 CTTTCATCGCGCGCTGGAGGAGCAAGGCTGTGGACATCATCTGCGCGCGCTGCGCA 1213
DB 1604 GNTTYATHGNGMNGTNGAYCAYCARAARGNGTNGAYATHTGCGNGAYGCTNATHCAYT 1663
QY 1214 AGATCTGCGCACCCCAAGGTGCAGATCGCATCTGCTGCGGCAAGCGCGCTGCTGCTGCT 1273
DB 1664 GGAH--GCGNGCARGAYGTNCARYTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 1720
QY 1274 AGAAGTGGTGAACGCGATCGGACCGACCAAGTACAAGSCCGCGCGCAAGGCGCTGCTGCT 1333
DB 1721 ARGAYATGTYMNGMNTTYGARWSNGARCAYSNGAYAAARNGMNGTNGGTTNGNT 1780
QY 1334 TCTCGCGCGCGCTGGCGCATGCTACCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 1393
DB 1781 TYWSNGTNCNYTNGCAYTNGNATHACNCGNGCNGAYATHTYNTYATGCGNWSNM 1840
QY 1394 GCTTCAGCGCGCTGCGCGCTGATCCAGCTGACCGCATGCTGCTGCTGCTGCTGCTGCTGCT 1453
DB 1841 GNTTYGARCCNTGCGGNTTNAAYCARYTNTAYGCTNGTNGTNGTNGTNGTNGTNGTNGTNG 1900
QY 1454 TAGCTTCACCGCGCGCTGCTGACACCGCTCAAGGAGGCGCTACCGGCTTCCATGCTG 1513
DB 1901 TNCAYGNGTNGGNGTNGMNGAYACNGTNGCNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 1960
QY 1514 GCGCGCTGAACCGCGCACAGCTGGACGAGGTGACGCGCGCGCTGCGCGCGCGCTGCGCG 1573

Db 738 CCCAAGGTGGACGTGACCTGCGGTGCGACGGCTACACCAACTACTCCCTCGAGACACT 679
Qy 1074 CTACGCGGCAAGCGCGCCGCAAGAGGCCCTGAGCCGAGCTGGGCTGCCTCTGGA 1133
Db 678 CGACGCTGGAAGAGCGCAGTGCAGAGCGGCCCTGAGCGGGAGCTGGGCTGGAAGTGG 619
Qy 1134 CCCAAGCGCCCTGTTGCGCTTCATCGCGCCCTGAGGAGCAGAAAGGTGTGGACAT 1193
Db 618 CGACGAGCTGCGCTGCTGCTGCTTCATCGGGCTCTGATGACAGAAAGGCGTGGACAT 559
Qy 1194 CATCTGCGCGCCCTGCGCAAGATCTGCGCACCCCAAGGTCAGATCGCATCTCTGG 1253
Db 558 CATCGGGGACGCGATGCGCGCTGATC---GCGGGGACGAGCTGCGACTGCTGATGCTGG 502
Qy 1254 TACCGCAAGCGCGCTAGCAGNAGTGTGACGCGCATCGGACCAAGTACAAGGCGG 1313
Db 501 CACCGGCGCGCGAGCTGGAAGAAATGCTGAGCACTTGGAGCGGAGCATCCCAACAA 442
Qy 1314 CGCAAGGCGGTGGTCAAGTTCCTGGCGCCCTGCGGCACATGCTCACCGCGCGGCCGA 1373
Db 441 GGTGCGCGGTGGTGGTTCCTGGTGCCTATGGCGCATCGCATCACGGCGCGGCCGA 382
Qy 1374 CTTATGCTGGTGGCTCGCGCTTCAGCCCTGCGGCTGATCCAGCTGCACGCCATGCA 1433
Db 381 CBTGCTGGTGGTGGCTTCGCGCTTCAGCCCTGCGGCTGAACCACTTACCGCGATGG 322
Qy 1434 CTACGGTACCGTGGCGGTGAGCTTCCACCGCGCGCTGGTGGACACCGCTCAAGGAGG 1493
Db 321 ATAGGCAACCGTCCCTGCTGTCACCGCTGGCGGCTCAGGACACCGTGGCGCGCTT 262
Qy 1494 CBTACCGCTTCCACATGGCGCCCTGAACCCCGACAAAGCTGGAGAGCTGACGCCGA 1553
Db 261 CGACCGCTTCAGGACGCGCGGTGCGGTGGACTTTGACCGYCGGAGGCCCAACAAGT 202
Qy 1554 CGCCCTGGCGCCACCGTGGCGCTGCCAGCGAGTGTTCGCGGCGCGCTACCCCGA 1613
Db 201 GATCGAGGCGCTCAGGCACTGCTCAGACAGTACCGGAACTACGAGAGAGCTGGAAGAG 142
Qy 1614 GATGGTGGCAACTGATCAGCAGGACCTGCTGCTGCTCAAGCCCGCCAGAGTGGGA 1673
Db 141 TCTCAGGCGCGGATGTCGAGACCTCAGCTGGGACACCGGCTGAGCTCTACGA 82
Qy 1674 GGCCTGCTGGAGAGGTGGTGTAC 1698
Db 81 GGACGCTCTGTCAAGGCCAAGTAC 57

RESULT 12
US-09-345-214-15
; Sequence 15, Application US/09345214
; Patent No. 6392120
; GENERAL INFORMATION:
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Broglie, Karen E.
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE
; FILE REFERENCE: BB-1147
; CURRENT APPLICATION NUMBER: US/09/345,214
; CURRENT FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: 060/094,436
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Zea mays
; US-09-345-214-15

Query Match 9.7%; Score 206.6; DB 4; Length 2019;
Best Local Similarity 55.0%; Pred. No. 9.4e-27;
Matches 476; Conservative 1; Mismatches 370; Indels 18; Gaps 3;

Qy 849 CAAGAAGATCAACTGCTGAAGGGTGGCATTATCGCCGCGGACAAAGCTGGTACTGTGTC 908
Db 887 CGAGCACGCCAACATCTTTGCGCGGGTCTGAAGATGGAGAGCCGGTGGTACTGTAG 946
Qy 909 GCCCAACTACGCGACCGAGATCGTCCGATCGCGCGGGTGTGAGCTGGACACCGT 968
Db 947 CCGCGGCTACCTGTGGAGCTGAAGACAGTGAAGCGGGCTGGGCTCCACGACATCAT 1006
Qy 969 C---ATCCCGCGCAAGGGCATTGAGGGCAITGTGAACGGCATGGACATGAGAGTGGAA 1025
Db 1007 CCGTCTAAGCATGGAAGATCAATGSCATCGTGAACGGCATGACACCGAGGAGTGA 1066
Qy 1026 CCCAAGACCGCAAGTTCCTGCTGCGCCCTAGACACCAAGACAGCG-----T 1073
Db 1067 CCCAAGTGGACCTGACCTCGCGTGGACGGCTACCAACTACTCCCTCGAGACACT 1126
Qy 1074 CTACGCGCGCAAGCGCCGCCAAGGAGGCCCTGAGCGCGAGCTGGGCTGCTGTGGA 1133
Db 1127 CGAGGCTGGAAGCGGAGTGAAGCGGCCCTGAGCGGGAGCTGGGCTTGAAGTGG 1186
Qy 1134 CCCACCGCGCCCTTCTTCGCTTCATCGCGCGCTGGAGGAGCAGAAAGGTGTGGACAT 1193
Db 1187 CGACGAGCTGCGCTGCTCGCTTCATCGGGCGTCTGGATGAGACAAAGGCGTGGACAT 1246
Qy 1194 CATCTGGCGCGCTGCCCAAGATCCTGCGCACCCCAAGTGCAGATCGCCATCTCTGGG 1253
Db 1247 CATCGGGAGCGCATGCGGTGGATC---GCGGGGACAGAGCTGCGACTGTGTGCTGG 1303
Qy 1254 TACCGCAAGCGCGCTTACGAGAAGCTGGTGAACGGCTCGGCACCAAGTACAAAGGCGCG 1313
Db 1304 CACGGCGCGCGACCTGGACGAATGCTGCAGCACTTGGAGGGGAGCATCCCAACAA 1363
Qy 1314 CGCAAGGCGGTGGTCAAGTTCCTGGCGCGCTGGCGCACATGCTACCGCGCGCGCGGA 1373
Db 1364 GGTGCGGGGTGGTGGTGGTTCCTGGTGGCTATGGCGCATCGCATCGCGCGCGCGCGGA 1423
Qy 1374 CTTCACTGCTGGTGGCTGCGCTGCGCTGCGACCGCTGATCCAGCTGCGACGCGCATGCA 1433
Db 1424 CGTCTGCTGATGCGCTGCGCGCTGCGAGCCCTGCGGCTGACCCAGCTCTACGCGATGGC 1483
Qy 1434 CTACGCTACCGTGCCTGCTGCTAGCTCCACCGCGCGCTGCTGCTGACACCGCTCAAGGAGG 1493
Db 1484 ATAGGCAACGCTCCCTGCTGGTGGACGCGCTGGGCGGCTCAGGACACCGTGGCGCTT 1543
Qy 1494 CGTACCGGCTTCCATGATGGCGCGCTGGAACCCCGACAAAGTGGACGAGGTGACGCGGA 1553
Db 1544 CGACCGCTTACGCGACCGCGGCTCGGCTGGACTTTTACCGYCGCGGCGCAACAAGCT 1603
Qy 1554 CGCCTGGCGCCACCGTGGCGCGTGGCAGCGAGTGTTCGGGCGCGCGCTACCCCGA 1613
Db 1604 GATGAGGCGCTCAGGCACCTGCCCTCGACACACTCGGNACTACGAGGAGCTGGAAGAG 1663
Qy 1614 GATGTTGCCCAACTCCATCAGCCAGGACCTGCTCTGTTGCCAAGCCCGCCAGAGTGGGA 1673
Db 1664 TCTCAGGCGCGGCGCATGTCGAGGACCTCAGCTGGGACCAACCGCGGTGAGCTCTACGA 1723
Qy 1674 GGGCGTGTGAGGAGGTGGTGTAC 1698
Db 1724 GGAGTCTCTGTCAAGGCCAAGTAC 1748

RESULT 13
US-08-941-445A-8
; Sequence 8, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle

```

: CITY: Boulder
: STATE: CO
: COUNTRY: US
: ZIP: 80303
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/941,445A
: FILING DATE: 30-SEP-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/026,855
: FILING DATE: 30-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Winner, Ellen P
: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 89-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ. ID. NO. 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2007 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: not relevant
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHEetical: NO
: ORIGINAL SOURCE:
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..2007
: US-08-941-445A-8

```

TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-572-951-2

Query Match 7.6%; Score 161; DB 1; Length 2085;
Best Local Similarity 53.2%; Pred. No. 4e-19;
Matches 444; Conservative 0; Mismatches 370; Indels 21; Gaps 4;

```
QY 870 GGTGGCATTATCGCGCCGACAAAGCTGTGACTGTGTCGCCCAACTACGCGACCCAGAT 929
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 985 GTGTGTCTGAAGATGGACAGCGGTGTGTGACTGTGACCGCGCGTACTGTGGAGCT 1044
QY 930 CGCTGCCGATCGCGCGCGGTGTGGAGCTGGACACCGTC---ATCCGCGCAAGGCGAT 986
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1045 GAAGACAGTGGAGGGGTGGGGCTCCACGACATCATCCGTTCTAAGACTGGAAGT 1104
QY 987 TGAGGCAATGTGAAGCGATGGACATTTGAGGAGTGGAAACCCCAAGCAGCAAGTTCTT 1046
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1105 CAATGCAATGTGAACGATCGACCAACAGGAGTGGAAACCCCAAGTGGACGTGCACCT 1164
QY 1047 GTCTGGCCCTAGGACCAAGAGCG-----TCTAGCGCGCAAGGCCCGCGC 1094
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1165 GCGTCCGAGCGGTACACCAACTACTCCCTCGAGACACTCGAGCTGGAAAGCGGAGTG 1224
QY 1095 CAAGGAGGCGCTCGAGCGCGAGCTGGGCTGCTGTGGACCCCAACCGCCCTGTTCCG 1154
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1225 CRAAGGCGGCGCTCGACGGGACGTGGGCTGGAGTGGCGGAGAGCTGCCGTGCTCGG 1284
QY 1155 CTTATCGGCGGCTGGAGGAGCAAGAGGTGTGGACATCATCTGGCGGCGCTGCCCAA 1214
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1285 CTTATCGGCGGCTGGATGGACAGAGGCGGTGGACATCATCGGGGACGCGATGCGGTG 1344
QY 1215 GATCTGGCCACCCCAAGTGCAGATCGCCATCTCTGGGTACCGGCAAGGCCCTACGA 1274
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1345 GATC---CGGGGCGAGAGTGCAGTGTGTGCTGGGACCGCGCCACCTGACCTGGA 1401
QY 1275 GAAGCTGGTGAACCGCATCGGACCAAGTACAAGGGCGCGCGCAAGGGCGTGTCAAGTT 1334
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1402 ACGAATGCTGCAGCACTTGGAGCGGAGCATCCCAACAAGGTGCGCGGTGGTTCGGTT 1461
QY 1335 CTGGGCGGCGCTGGCGACATGCTACCGCGCGCGCGGACTTCATGCTGTGTGCGCTCGCG 1394
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1462 CTGGTCTTAATGGTCATCGCATCACGCGCGCGCGGAGCGTGTGTGTATGCTGCTCCG 1521
QY 1395 CTTGAGCGCTGGCGCTGATCAGCTGCACGCTACGCTACGCTACGCTGCGCTGCGTGT 1454
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1522 CTTTCG---CGGGGCGGTGAACAGCTTACGCGATGGCATCGGACCGCTGCTGTGT 1578
QY 1455 AGCTCCACCGCGGCTGTGTGACACCGCTCAAGGAGGGGCTCACCGGCTTCCACATGGG 1514
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1579 GCAGCGGTGGCGGCTCAGGAGACCGGTGGCGGCTTCGACCGCTTCGCGAGCGCGG 1638
QY 1515 CGCCCTGAACCCGACAAAGTGAAGAGGTGAGCGCGGAGCGCTGCGCGCGCACCGTGG 1574
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1639 GCTCGGTTGACTTTTGACCGCGCGCGAGGCGCAACAGCTGATCGAGGTGTGCGCCACTG 1698
QY 1575 CCGTGCAGCGAGGTGTTTCGGCGGCGCGCTACCCCGAGATGTGGCCCAACTGCATCAG 1634
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1699 CCTGACACGTACCGAATACGAGGAGAGCTGGAGAGAGTCTCCAGGCGCGGCACTGTC 1758
QY 1635 CCAGGACCTTCTCTGTCAGCGCGCGCGCGCGAGAGTGGAGGGCGCTGCTGGAGGAG 1689
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1759 GCAGAACCTCAGCTGGGAGCACCGCGGTGAGCTCTACGAGGAGCGTCTCTTGTCAAG 1813
```

RESULT 15

US-09-196-390-1
; Sequence 1, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina

```
APPLICANT: Lorz, Horst
APPLICANT: Luticke, Stephanie
APPLICANT: Walter, Lennart
APPLICANT: Froberg, Claus
APPLICANT: Kossmann, Jens
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
TITLE OF INVENTION: SYNTHESIS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: James F. Haley, Jr., c/o Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/196,390
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 196 21 588.9
FILING DATE: 29-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 196 36 917.7
FILING DATE: 11-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP97/02793
FILING DATE: 28-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Haley, Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: AGREVO-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2239 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Triticum aestivum L.
STRAIN: cv. Florida
HAPLOTYPE: ca. 21 d Caryopses
IMMEDIATE SOURCE:
LIBRARY: cDNA library in pBluescript sk (-)
CLONE: TASSS
FEATURE:
NAME/KEY: CDS
LOCATION: 3..2017
US-09-196-390-1
```

Query Match 5.6%; Score 118.2; DB 4; Length 2239;
Best Local Similarity 52.4%; Pred. No. 5.9e-12;
Matches 333; Conservative 0; Mismatches 293; Indels 9; Gaps 3;

```
QY 857 TCACTGCTGAAGGTGGCATTATTCGCCCGCGACAAAGCTGCTGCTGCGCCAACT 916
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 823 TTAACCTTTTGAAGAGGACGACTTGTGACAGAGATCGGATGTGACCGCTCAGTCAAGGTT 882
QY 917 ACGCCACCGAGATC---GCTCCGATACCGCGCGCGGTGTGAGCTGGACACCGCTCAATCC 973
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 883 ATTCAATGGAGGTCACAACTGCTGAAGGTGGACAGGGCGCTCAATGAGCTCTTAAGCTCCC 942
```

QY	974	GGCGCAAGGGCAATTGAGGCATTTGTGAACGGCATGGACATTCAGAGTGGAAACCCCAAGA	1033
Db	943	GAAGAAAGTGTTATGAATGGAAATTGTAATGGAAATTGACATTAAATGATTTGGAAACCCCA	1002
QY	1034	CCGACAGATTCCTGTCTGCGGCCCTTACGACCAAGACAGCGTCTACGCGGCAAGCCGCGG	1093
Db	1003	CAGACAGTGTCTCCCTCATCATTAATCTGTGATGAGCTCT--CTGGAAGGCCAAAT	1059
QY	1094	CCAAGGAGGCCCTGCAGGCGGAGCTGGCCCTGCCTGTGGACCCACCGCCCCCTGTTCG	1153
Db	1060	GTAAGCTGAATTCGAGAGAGTGTGGGTTTACCTGTAAAGGAGATGTTCTCTGTGATTG	1119
QY	1154	CCTTCATCGGCCCTGTGAGGAGCAAGAGGTGTGGACATCATCTGGCGCCCTGGCCCA	1213
Db	1120	GCCTTATTGGAAGACTGGATTACCAGAAAGGCATGATCTCATTAATAGCCATTCACG	1179
QY	1214	AGATCCTGGCCACCCCAAGGTGCGAGATGCGCATCCTTGGGTACCGCAAGCGCCCTACG	1273
Db	1180	AGCTCATG--AGGAGGAGCTGCAATTTGTATGCTTGGATCTGGGATCCCAATTTTG	1236
QY	1274	AGAACTGTTGACGCCATCGSCACCAAGTACAAGGCCGCCGCAAGGGGTGGTCAAGT	1333
Db	1237	AAGCTGGATGAGATCTACCGAGTCGAGTTTACAAGGATAAATTCCTGGATGGGTTGGAT	1296
QY	1334	TCTCGGCGCCCTGGCGCACATGCTCACGGCGGGCGGAGCTTCATGCTGTGCCCTCGC	1393
Db	1297	TTAGTGTTCAGTTTCCACACAATAACTGCAAGTTCGAGTTGGATATATGTTAATGCGCATCGA	1356
QY	1394	GCTTCGAGCCCTGCGGCTGTATCCAGCTGCACGCCATCACTACGGTACCGTGCCCGTGG	1453
Db	1357	GATTTGAACCTTGGGCTCTTAATCAGCTATATGTCTGCAATATGTTACAGTTCCTGTAG	1416
QY	1454	TAGCCTCCACGGCGCCCTGTCGACACCGCTAAG	1488
Db	1417	TTGATGGAATGGGGCCCTCCGAGACACAGTTCGAG	1451

Search completed: June 3, 2003, 15:30:36
Job time : 81.0928 secs

This Page Blank (uspr)

GenCore version 5.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 21:49:59 ; Search time 192.814 Seconds
(without alignments)
15416.986 Million cell updates/sec

Title: US-09-980-771A-2

Perfect score: 2124

Sequence: 1 atggctgtgctctaccag.....ccgccaccgcgaagtcgcgc 2124

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	478	22.5	2267	9	US-09-961-077-25
2	314.8	14.8	2399	9	US-10-138-075-1
3	277.8	13.1	2179	9	US-10-138-075-3
4	218.2	10.3	2825	10	US-09-952-677-5
5	118.2	5.6	2239	10	US-09-952-677-1
6	102	4.8	2712	10	US-09-748-033-4
7	94.2	4.4	790	10	US-09-966-881-44
8	86.6	4.1	1266	9	US-10-145-415-98
9	84.2	4.0	2307	9	US-09-893-519A-87
10	83.2	3.9	1427	10	US-09-974-300-757
11	82.2	3.9	805	10	US-09-966-881-43
12	81	3.8	12441	9	US-09-988-384B-3
13	81	3.8	13613	9	US-09-860-846-3
14	81	3.8	13613	9	US-09-836-821-3
15	81	3.8	13613	10	US-09-861-289-3
16	76.4	3.6	813	10	US-09-966-881-45
17	76	3.6	1318	9	US-09-934-900-15
18	75.4	3.5	446	10	US-09-864-761-20699
19	75.2	3.5	1896	9	US-10-124-880-15

20	74.8	3.5	1407	10	US-09-815-242-4089	Sequence 4089, Ap
21	74.8	3.5	15872	9	US-09-860-846-1	Sequence 1, Appli
22	74.8	3.5	15872	9	US-09-988-384B-1	Sequence 1, Appli
23	74.8	3.5	15872	9	US-09-836-821-1	Sequence 1, Appli
24	74.8	3.5	15872	10	US-09-861-289-1	Sequence 1, Appli
25	74.6	3.5	1266	9	US-10-145-415-104	Sequence 104, App
26	74.2	3.5	1770	9	US-10-100-957A-1	Sequence 1, Appli
27	73.6	3.5	88421	9	US-09-976-059-1	Sequence 1, Appli
28	72.4	3.4	3468	9	US-09-988-462-2	Sequence 2, Appli
29	72.2	3.4	299	10	US-09-294-093B-3520	Sequence 3520, Ap
30	72.2	3.4	824	10	US-09-894-633A-83	Sequence 83, Appl
31	71.4	3.4	3624	9	US-09-988-462-6	Sequence 6, Appli
32	71.4	3.4	12733	9	US-10-032-393-47	Sequence 47, Appli
33	71.4	3.4	12739	9	US-10-032-393-8	Sequence 8, Appli
34	71.2	3.4	1182	9	US-09-364-847-1	Sequence 1, Appli
35	71.2	3.4	1926	9	US-09-364-847-9	Sequence 9, Appli
36	71.2	3.4	1926	9	US-09-364-847-18	Sequence 18, Appl
37	70.8	3.3	1415	9	US-09-934-900-11	Sequence 11, Appl
38	70.6	3.3	1635	10	US-09-864-761-20241	Sequence 20241, A
39	70.6	3.3	1973	10	US-09-864-761-3471	Sequence 3471, Ap
40	69.8	3.3	1075	10	US-09-864-761-19241	Sequence 19241, A
41	69.8	3.3	1403	10	US-09-864-761-2513	Sequence 2513, Ap
42	69.8	3.3	2010	12	US-10-032-717-9	Sequence 9, Appli
43	69.8	3.3	4889	9	US-09-860-846-34	Sequence 34, Appl
44	69.8	3.3	4889	9	US-09-988-384B-34	Sequence 34, Appl
45	69.8	3.3	4889	9	US-09-836-821-34	Sequence 34, Appl

ALIGNMENTS

RESULI 1

US-09-961-077-25

; Sequence 25, Application US/09961077

; Publication No. US20030014775A1

; GENERAL INFORMATION:

; APPLICANT: Zwick, Michael G.

; Edington, Brent E.

; McSwiggen, James A.

; Merlo, Patricia Ann Owens

; Guo, Lining

; Skokut, Thomas A.

; Young, Scott A.

; Folkerts, Otto

; Merlo, Donald J.

TITLE OF INVENTION: COMPOSITION AND METHODS FOR
MODULATION OF GENE EXPRESSION
IN PLANTS

NUMBER OF SEQUENCES: 1263

CORRESPONDENCE ADDRESS:

ADDRESSER: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Suite 4700

STATE: Los Angeles

COUNTRY: California

ZIP: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/961,077

FILING DATE: 21-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/679,645

FILING DATE: July 12, 1996

APPLICATION NUMBER: 60/001,135

FILING DATE: July 13, 1995

APPLICATION NUMBER: 08/300,726

[illegible]

QY	1199	TGCGCGCCCTGCCCCAGATCTTGGCCACCCCAAGTGCAGATCGCATCTCTGGTACCG	1358
Db	935	AGCGGGTATTGTCATGAACCTGTTTGAAGAAGACACATGCAGCTGATCGTTTATAGAACCG	994
QY	1259	GCAAGCGCGCTACGAGAAGCTGGTGAACGCCATCGGCACCAAGTACAAGGGCGCGGCCA	1318
Db	995	GAGAAGCGGNAATTGGAATTAACCTCCGTATATGAGAGCATGCTTTCCGGACAGGTGCA	1054
QY	1319	AGCGGTGGTTCATAGTTCTTCGGCGCCCTGGCGCACATGCTACCGCGCGCGCACTTCA	1378
Db	1055	AGCGGTATATCGGTTTTCATGAACCGCTACGCCCAAGATTTATGCGCGAGCGATCTTT	1114
QY	1379	TGCTGTGCGCTCGCGCTTCGAGCCCTGCGGCTGATCAGCTGCAGCCCATGCCTAACG	1438
Db	1115	TCTGATGCGCTTCAAAATTCGAGCGCTGGCGTCTTGACAGCTGATGCGCTCAGGTACG	1174
QY	1439	GTACCGTGGCGGTGGTAGCTCCACCGCGCGCGCTGGTCCGACACCGTCA	1486
Db	1175	GCSCGTGCGGTTGTGTCAGGGAACAGCGCGGCTTCATGATACGGTGA	1222

RESULT 11
 US-09-966-881-43
 ; Sequence 43, Application US/09966881
 ; Patent No. US20020120960A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Seymour, Graham
 ; Bird, Colin
 ; Medina-Suarez, Rosybel
 ; TITLE OF INVENTION: Genetic control of Fruit Ripening
 ; NUMBER OF SEQUENCES: 57
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Zeneca Ag Products Inc.
 ; STREET: 1800 Concord Pike
 ; CITY: Wilmington
 ; STATE: DE
 ; COUNTRY: USA
 ; ZIP: 19850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/966.881
 ; FILING DATE: 28-Sep-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/242.860
 ; FILING DATE: 29-Mar-1999
 ; APPLICATION NUMBER: GB 9618862.8
 ; FILING DATE: 10-SEP-1996
 ; APPLICATION NUMBER: GB 9708366.1
 ; FILING DATE: 25-APR-1997
 ; APPLICATION NUMBER: PCB/GH97/02424
 ; FILING DATE: 08-SEP-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hohenschutz, Liza D.
 ; REGISTRATION NUMBER: 33,712
 ; REFERENCE/DOCKET NUMBER: SEE 50183/UST
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (302) 886-1699
 ; INFORMATION FOR SEQ ID NO: 43:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 805 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; IMMEDIATE SOURCE:
 ; CLONE: U-D66
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 43:

Db 11421 TTCCGAGGTATCATGACGCGAACCAGGCAACACCGCCCTACCGGAGCACCCTCCGG 11362
QY 1208 TGCCCAAGATCTTGGCCACCCCAAGGTGAGATGCGCATCTCTGGGTACCGGAAGCCG 1267
Db 11361 ACCTCCCGCGCTCTCTGTCGCGGACACGACGCGCCACGCGCTCAACAACACCACTACG 11302
QY 1268 CTTACGAGAGCTGGTGAACGCGCATCGGACCAAGTACAAGGCGCGCCGCAAGGCGTGG 1327
Db 11301 TGATCTCGAGATCGAGAGGCGCACCGGCGCATCCACCGGACCTCGTATGAGGTCC 11242
QY 1328 TCAAGTCTGCGGCGCCCTTGGCGACATGCTCAACGCGCGCGCGGCTTCAATGCTGTGC 1387
Db 11241 TGAAGCGGAAGCGTGCACACCGCGCTTCTTCTGCGGCGCTGCCAGAGTGGAGC 11182
QY 1388 CTTGCGGCTTCGAGCGCTCGGGGCTGATCCAGCTGACGCGCATGCACTACGTTACCGTGC 1447
Db 11181 CGTACCGCGGCGGCGCGACGCGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGTGC 11122
QY 1448 CGGTGTAGCTCTCACCGCGGCGCTGTGCGACATGCTCAACGCGCGCGCGGCTTCAATGCTGTGC 1307
Db 11121 TGTCCCTGGACCGGCGACCGCATCGGCGAGCAGCAGATCCCGCGGCTGCGCGACCTGC 11062
QY 1508 ACATGGGCGCTTGAACCGCGCAAGCTGGACGAGCTGACGCGGCGCGCGCTGCGCGGCA 1567
Db 11061 TGGTCTCTGCGGCGCGCGCGCAACTGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 11002
QY 1568 CGTGGCGGCTGCCAGCGAGGTGTTTGGGCGCGCGCGCTACCGCGGAGTGTGGCGCAACT 1627
Db 11001 CCGCGCTCGGCG 10952
QY 1628 GCATCAGCAGGAGCTCTCTGTGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1587
Db 10951 TGACCG 10892
QY 1688 AGGTGTGTACGCGCAAGGCGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1747
Db 10891 ATCTGGGCGCGCGGTCCAGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10832
QY 1748 TTGCGGAGAGATCCCGCGGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1807
Db 10831 CGACAGGCGCGGAGCG 10772
QY 1808 CGGTGTCCCGCTCGGTGAGGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1867
Db 10771 GCGGCTTACCGCGGTGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10712
QY 1868 CGGCATGGGCGGTGGCG 1927
Db 10711 GCGGCTTCCCGCTGTTGCGGCGAGTCCCGGAGCTCCCGCGCGCGCGCGCGCGCG 10652
QY 1928 CCAAGGTACACCTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1987
Db 10651 CGCGGCGGCTTGTGAAGAACACCTCTCCCGCTCGAAGCAGCGCGCGCGCGCGCG 10592
QY 1988 TCAAGCTGCGCGGTGAGGCTTCCACACCTCGACCTCGGAGAACGCGCGCTGCTTCAAG 2047
Db 10591 CGCGCTCGCGGAGCGCGGTCTCCCGTACAGGCTCGCGCTTACCGCGCGCGCGCG 10532
QY 2048 GCA 2050
Db 10531 GCA 10529

RESULT 13

US-09-860-846-3/c
; Sequence 3, Application US/09860846
; Patent No. US2002016472A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin

; FILE REFERENCE: 600.438051
; CURRENT APPLICATION NUMBER: US/09/860.846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 13613
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-860-846-3

Query Match 3.8%; Score 81; DB 9; Length 13613;
Best Local Similarity 42.7%; Pred. No. 1.3e-09;
Matches 488; Conservative 0; Mismatches 645; Indels 10; Gaps 1;

QY 908 CGCCCAACTACGCGACCGAGATGCTGCCGATGCCGCGGCGGTGTGAGCTGGACACCG 967
Db 12348 CGGTGACG 12389
QY 968 TCATCCGCGCGCAAGGGCATTTGAGGGCATTTGACGCGCATGGACATTTGAGAGTGGAAC 1027
Db 12288 CCAAGGCGGTCAACGCGCTTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12229
QY 1028 CCAAGACCGCAAGATTTCTGTCTGCGCGCTTACGACCAAGACAGCGCTTACGCGGCAAG 1087
Db 12228 CCGGATCCGCGCGCTTCCACAACTTCGCGCTTCGACCTGCCGCGCGCGCGCGCG 12169
QY 1088 CCGCGCGCAAGGAGCGCTCGAGGCGAGCTGGGCGCTTGGAGCGCGCGCGCGCGCG 1147
Db 12168 GGACCAACGCAAGATGAGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12109
QY 1148 TGTTCGCTTCATCGCGCGCTTGAGGAGCAGAGGCTGTGGACATCATCTTGGCGCGCG 1207
Db 12108 TTCCGAGGTCTACGAGCGGAAACCGGCGCAACCGCGCGCGCGCGCGCGCGCGCG 12049
QY 1208 TGCCCAAGATCTCTGGCGCACCCCAAGGTGAGATCGCATCTCTGGTACCGCAAGCGCG 1267
Db 12048 ACCTCCGCGCGTCTCTGTCGCGCGACCGACCGCGCGCGCGCGCGCGCGCGCGCG 11989
QY 1268 CTTACGAGAGCTGGTGAACGCGCATCGGACCAAGTACAAGGCGCGCGCGCGCGCG 1327
Db 11988 TGATCTGAGATCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11929
QY 1328 TCAAGTCTCGCGCGCGCTGGCGCATGCTCAACGCGCGCGCGCGCGCGCGCGCGCG 1387
Db 11928 TGAAGCGCGAAGCGGTGCACACCGCGCGCTTCTTCTGCGGCGCTGCCAGAGCTGGAGC 11869
QY 1388 CTTGCGGCTTCGAGCGCGCTGGGCGCTGATCCAGCTGACGCGCATGCACTACGTTACG 1447
Db 11868 CGTACCGCGGCG 11809
QY 1448 CCGTGTAGCTCTCACCGGCGCGCTGTGCACACCGCTCAAGGAGGCGGTACCGGTTCC 1507
Db 11808 TGTCCCTGCG 11749
QY 1508 ACATGGGCGCGCTTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1567
Db 11748 TGGCTCTCTGCG 11689
QY 1568 CCGTGGCGGTGCCAGCGAGGTGTTTGGGCGGCGCGCTTACCGCGAGATGGTGGCGCACT 1627
Db 11688 CCGCGCTCG 11639
QY 1628 GCATCAGCAGGAGCTGCTCTGTGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1687
Db 11638 TGAACCG 11579
QY 1688 AGGTGTGTACGCGAAGGCGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1747
Db 11578 ATCTGGGCGCGGTTCCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11519

1748 TTGCGGAGAGATCCCGGCGACCTGCGCGCGCTGCTTCTAGCGCCCAACACCCCTGAAGC 1807
11518 CCGACGAGCGGAGAACGACCGCGCCACCTGCTCGGCGTCCGCGTACGSCAACA 11459
1808 CCGTGTCCGCTCGTGGAGGCGACGCGCGCGCGCGCCCAAGTCTGGCAGCACCAGCC 1867
11458 GCGCTTACACCCGCTGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11399
1868 CCGCATGGCGGCTGGCG 1927
11398 GCGCGCTCGCTGCTGGCG 11339
1928 CCAAGGTGACCACTTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1987
11338 CCGCGCGGCGCTACTGGAAGAACACCTGCTCCCGCTCGAACAGCGCGCGCTTCGAGC 11279
1988 TCAAGTGGCGGCTGAGCGCTCCACCACTCGACCTCGAGACGCGCGCTGCCTCCAAGC 2047
11278 GCGCGCTCGCGAGGAGCGCGCTTCTCCGCTACAGCGTCCGCGCTTACCGCGCGCGCG 11219
2048 GCA 2050
11218 GCA 11216

RESULT 14
US-09-836-821-3/c
; Sequence 3, Application US/09836821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, Y.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 13613
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-836-821-3

Query Match 3.8%; Score 81; DB 9; Length 13613;
Best Local Similarity 42.7%; Pred. No. 1.3e-09;
Matches 488; Conservative 0; Mismatches 645; Indels 10; Gaps 1;

908 CCGCCCAACTACCGACCGAGATCGCTCGCGATCGCGCGCGCGCGTGTGGAGTGGACACCG 967
12348 CGGTGACG 12289
968 TCATCCG 1027
12288 CCAAGGCGCTCAACGCGCTTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12229
1028 CCAAGACCGACAGTTCCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1087
12228 CCGCGATCG 12169
1088 CCG 1147
12168 GGACCAACCGCAAGATGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12109
1148 TGTTCGCGCTTCATCG 1207
12108 TTCCGAGGTCTATCGACGGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12049
1208 TGCCCCAGATCTTGGCCACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1267

12048 ACCTCCCGCGCGCTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11989
1268 CCTACGAGAAGCTGCTGAAGCGCATCGCGCAACAGTACAAAGCGCGCGCGCGCGCGCG 1327
11988 TGATGCTGAGATCGACGAGCGCACCGCGCATCCACCGCACTCGTCTGATGGAGTCC 11929
1328 TCAAGTTCCTGCG 1387
11928 TGAAGCGCGAAGCGCTGACACCGCGCGCTACTTCTCGCGCGCGCTGCCACGAGCTGG 11869
1388 CCGTCCGCTTCGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1447
11868 CGTACCG 11809
1448 CCGTGTAGCTGCG 1507
11808 TGTCCTCG 11749
1508 ACATGGGCGCGCTGAACCG 1567
11748 TGCGTCTCTGCG 11689
1568 CCGTGGCGCGTCCGACGAGGCTTTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1627
11688 CCGCGCTCG 11639
1628 GCATCAGCGAGGACCTGCTCTGCTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1687
11638 TGACG 11579
1688 AGTGTGTACGCGAAGCG 1747
11578 ATCTGGGCG 11519
1748 TTCCGCGAAGATCCCG 1807
11518 CCGACGAGCG 11459
1808 CCGTGTCCG 1867
11458 GCGCGCTTACCG 11399
1868 CCGCGATGGCG 1927
11398 GCGCGCTCTCTGCG 11339
1928 CCAAGGTGACCG 1987
11338 CCG 11279
1988 TCAAGTTCG 2047
11278 GCGCGCTCGCGAGGAGCGCGCTTCTCCGCTACAGCGTCCGCGCTTACCGCGCGCGCG 11219
2048 GCA 2050
11218 GCA 11216

RESULT 15
US-09-861-289-3/c
; Sequence 3, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18

This Page Blank (uspto)

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, .rnpm and .rnpn

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, .rapm and .rapn

The Pending database search results should not be left in the case because they contain data that is confidential.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 21:50:00 ; Search time 3310.08 Seconds
(without alignments)
16133.327 Million cell updates/sec

Title: US-09-980-771A-2

Perfect score: 2124

Sequence: 1 atgtgtgttcctaccag.....cgccaccgcaagtcgcgc 2124

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_Main.*
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2124	100.0	2124	37	US-09-980-771A-2
2	2124	100.0	3117	37	US-09-980-771A-1
3	1953	91.9	1953	37	US-09-980-771A-4
4	1593	75.0	1593	37	US-09-980-771A-8
5	1314	61.9	1314	37	US-09-980-771A-6
6	533.8	25.1	2058	3	US-07-771-351-4
7	527.4	24.8	2201	3	US-07-771-351-5
8	523.8	24.7	2279	3	US-07-771-351-3
9	478	22.5	1818	29	US-09-731-166-3
10	478	22.5	2267	7	US-08-300-726-1
11	478	22.5	2267	7	US-08-300-726A-1
12	478	22.5	2267	36	US-09-961-077-25
13	472.8	22.3	1915	18	US-09-402-254-46
14	449.2	21.1	2566	25	US-09-654-617-385561
15	449.2	21.1	2566	27	US-09-684-016-385561
16	447.6	21.1	2542	24	US-09-625-406-6
17	323.2	15.2	1160	3	US-07-771-351-6
18	314.8	14.8	2399	40	US-10-138-075-1
19	314.8	14.8	2399	72	US-60-288-315-1
20	305.2	14.4	2098	42	US-10-219-999-9995
21	305.2	14.4	2284	25	US-09-654-617-258750

22 305.2 14.4 2284 27 US-09-684-016-258750
23 300.6 14.2 1416 42 US-10-219-999-7933
24 300.6 14.2 1416 75 US-60-312-544-3778
25 298 14.0 1406 42 US-10-219-999-7327
26 298 14.0 1406 75 US-60-312-544-3248
27 298 14.0 1409 42 US-10-219-999-13187
28 298 14.0 2115 19 US-09-513-996A-779
29 298 14.0 2115 22 US-09-565-309A-59507
30 298 14.0 2115 22 US-09-595-328C-273
31 297 14.0 2115 22 US-09-565-309A-53899
32 296.2 13.9 1981 25 US-09-654-617-190692
33 296.2 13.9 1981 27 US-09-684-016-190692
34 281.2 13.2 2067 38 US-10-044-543-21
35 277.8 13.1 2179 40 US-10-138-075-3
36 277.8 13.1 2179 72 US-60-288-315-3
37 277.8 13.1 2353 25 US-09-654-617-274
38 277.8 13.1 2353 27 US-09-684-016-274
39 266.6 12.6 2274 38 US-10-044-543-17
40 259.6 12.2 2176 38 US-10-044-543-13
41 255.4 12.0 2202 38 US-10-044-543-1
42 253 11.9 836 25 US-09-654-617-259308
43 253 11.9 836 27 US-09-684-016-259308
44 221 10.4 2842 38 US-10-018-418-3
45 220.8 10.4 1295 25 US-09-654-617-270820

ALIGNMENTS

RESULT 1

US-09-980-771a-2

; Sequence 2, Application US/09980771a

; GENERAL INFORMATION:

; APPLICANT: CNRS

; TITLE OF INVENTION: STARCH GRANULES CONTAINING A RECOMBINANT POLYPEPTIDE OF INTEREST,

; FILE REFERENCE: METHOD OF OBTAINING THEM AND THEIR USES

; CURRENT APPLICATION NUMBER: US/09/980,771a

; CURRENT FILING DATE: 2001-11-15

; PRIOR APPLICATION NUMBER: FR 99/06494

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 2124

; TYPE: DNA

; ORGANISM: artificial sequence

; FEATURE:

; OTHER INFORMATION: fragment of the complete sequence of cDNA coding for the GBSSI of

; OTHER INFORMATION: Chlamydomonas reinhardtii

; NAME/KEY: CDS

; LOCATION: (1)..(2124)

; OTHER INFORMATION:

US-09-980-771a-2

Query Match

Best Local Similarity 100.0%; Score 2124; DB 37; Length 2124;

Matches 2124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTGTGCCTTACCAGCGCCCGCAGCAGCGCGTCTATCGTCAATATGCGCG 60

Db 1 ATGGCTGTGCCTTACCAGCGCCCGCAGCAGCGCGTCTATCGTCAATATGCGCG 60

Qy 61 TCGTTCGGTGTCAAGAGACCGGACACAGCTGCTCGGTGAGCTTGTCTGGCTCGGCA 120

Db 61 TCGTTCGGTGTCAAGAGACCGGACACAGCTGCTCGGTGAGCTTGTCTGGCTCGGCA 120

Qy 121 CGCAAGTCCACTCGCTCGCTGTTACTGGTGCCTGCTGCTGCTGCTGCTGCTGCTG 180

Db 121 CGCAAGTCCACTCGCTCGCTGTTACTGGTGCCTGCTGCTGCTGCTGCTGCTGCTG 180

Qy 181 ATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

Db 181 ATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

181 ATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
241 ACTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
241 ACTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
301 CGCTAGGACCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
301 CGCTAGGACCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
361 AAGTTCGGTACTTCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT 420
361 AAGTTCGGTACTTCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT 420
421 TGGTTCCTGGCCAAAGTCTGGGGCAAGACCGGCTTCAAGCTGTACGGCCCGCT 480
421 TGGTTCCTGGCCAAAGTCTGGGGCAAGACCGGCTTCAAGCTGTACGGCCCGCT 480
481 GCTGACTACCTGACCAACCAAGGCTTTCGCGCTTCTGCAAGCGCGCTATGAG 540
481 GCTGACTACCTGACCAACCAAGGCTTTCGCGCTTCTGCAAGCGCGCTATGAG 540
541 GCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
541 GCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
601 TCCGCCCTGGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
601 TCCGCCCTGGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
661 GCCAAGTCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
661 GCCAAGTCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
721 TTCAGGACACGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
721 TTCAGGACACGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
781 GCCAAGTTCACCTGAGGCGCCACCCATGAGGAGGAGGAGGAGGAGGAGGAG 840
781 GCCAAGTTCACCTGAGGCGCCACCCATGAGGAGGAGGAGGAGGAGGAGGAG 840
841 AAGACCTACAGAAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
841 AAGACCTACAGAAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
901 ACTGTGCGCCCACTAGCGACCGAGATGCTGCTGCTGCTGCTGCTGCTGCTG 960
901 ACTGTGCGCCCACTAGCGACCGAGATGCTGCTGCTGCTGCTGCTGCTGCTG 960
961 GACACCGTCTACCGCGCCCAAGGCGATTGAGGCGATTGTAACGCGCATGAGG 1020
961 GACACCGTCTACCGCGCCCAAGGCGATTGAGGCGATTGTAACGCGCATGAGG 1020
1021 TGGAAACCCCAAGACCGGACGATTCCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
1021 TGGAAACCCCAAGACCGGACGATTCCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
1081 GGCAAGCGCGCGCCCAAGGCGCTGCGAGCTGGAGCTGGGCTGGGCTGGGCT 1140
1081 GGCAAGCGCGCGCGCCCAAGGCGCTGCGAGCTGGGCTGGGCTGGGCTGGGCT 1140
1141 GCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
1141 GCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
1201 GCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
1201 GCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
1261 AAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
1261 AAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

QY 1321 GCGTGGTCAAGTCTCGCGCGCCCTCGGCGACATGCTCAACCGCGCGCCGACATTCATG 1380
Db 1321 GCGTGGTCAAGTCTCGCGCGCCCTCGGCGACATGCTCAACCGCGCGCCGACATTCATG 1380
QY 1381 CTGGTGGCCCTCGCGCTCGAGCCCTCGGCGCTGATCCAGCTGACGCCATGACATCGGT 1440
Db 1381 CTGGTGGCCCTCGCGCTCGAGCCCTCGGCGCTGATCCAGCTGACGCCATGACATCGGT 1440
QY 1441 ACCGTGCCGTGGTACCTCCACCGCGCGCCCTCGGCTGACACCGTCAAGGAGGCGTCACC 1500
Db 1441 ACCGTGCCGTGGTACCTCCACCGCGCGCCCTCGGCTGACACCGTCAAGGAGGCGTCACC 1500
QY 1501 GCTTCCACATGGGCGCCCTGTAACCCCGACACAGCTGGAGGCTGACGCCGACGCCCTG 1560
Db 1501 GCTTCCACATGGGCGCCCTGTAACCCCGACACAGCTGGAGGCTGACGCCGACGCCCTG 1560
QY 1561 GCGGCCACGTGCGCGTCCAGCGAGGCTGTTGCGGGCGCGCTACCCCGAGATGGT 1620
Db 1561 GCGGCCACGTGCGCGTCCAGCGAGGCTGTTGCGGGCGCGCTACCCCGAGATGGT 1620
QY 1621 GCCAACTGATCAGCCAGGACCTGCTCTGCTCCAGCCCGCCAGAGTGGAGGCGCTG 1680
Db 1621 GCCAACTGATCAGCCAGGACCTGCTCTGCTCCAGCCCGCCAGAGTGGAGGCGCTG 1680
QY 1681 CTGGAGGAGTGGTGTACGGGAAGGGCGCGCTGGCCACCGCCAAAGAGGAGATCAAG 1740
Db 1681 CTGGAGGAGTGGTGTACGGGAAGGGCGCGCTGGCCACCGCCAAAGAGGAGATCAAG 1740
QY 1741 GTGCCCGTTCGCGAGAAGATCCCGCGACCTGCGCGCGCTGCTACGCCCCCAACACC 1800
Db 1741 GTGCCCGTTCGCGAGAAGATCCCGCGACCTGCGCGCGCTGCTACGCCCCCAACACC 1800
QY 1801 CTGAAGCCCGTGTCCGCTCGTGGAGGCAACGGCGCGCGCGCCGCGCCAAAGTTCGCAAC 1860
Db 1801 CTGAAGCCCGTGTCCGCTCGTGGAGGCAACGGCGCGCGCGCCGCGCCAAAGTTCGCAAC 1860
QY 1861 ACCGCCCGCATGGCGCGTGGCGGAGACACCCCTCGGGCCCTCGCGCGCGCGCC 1920
Db 1861 ACCGCCCGCATGGCGCGTGGCGGAGACACCCCTCGGGCCCTCGCGCGCGCGCC 1920
QY 1921 GCCACCCCAAGTACACCTACAAGCCCGCGCTGCGCGCCACCGCCCAAGCCCAAGACC 1980
Db 1921 GCCACCCCAAGTACACCTACAAGCCCGCGCTGCGCGCCACCGCCCAAGCCCAAGACC 1980
QY 1981 GCTGGCCTCAAGTGGCGGTGAGGCTCCACACCTCGACCTCGGAGAACGCGCTGCC 2040
Db 1981 GCTGGCCTCAAGTGGCGGTGAGGCTCCACACCTCGACCTCGGAGAACGCGCTGCC 2040
QY 2041 TCCACGGCAACGGCAACGGTGGCTCGGCGCTCGCAAGACCTCGGCTGCCAAGCCCTGGTC 2100
Db 2041 TCCACGGCAACGGCAACGGTGGCTCGGCGCTCGCAAGACCTCGGCTGCCAAGCCCTGGTC 2100
QY 2101 TCCGCGCCACCGCAAGTCCGCG 2124
Db 2101 TCCGCGCCACCGCAAGTCCGCG 2124

RESULT 2
US-09-980-771a-1
; Sequence 1, Application US/09980771A
; GENERAL INFORMATION:
; APPLICANT: CNRS
; TITLE OF INVENTION: STARCH GRANULES CONTAINING A RECOMBINANT POLYPEPTIDE OF INTEREST,
; FILE OF INVENTION: METHOD OF OBTAINING THEM AND THEIR USES
; FILE REFERENCE: WO 99 AB CNR AMYL
; CURRENT APPLICATION NUMBER: US/09/980, 771A
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: FR 99/06494
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1

; LENGTH: 3117
; TYPE: DNA
; ORGANISM: Chlamydomonas reinhardtii
US-09-980-771a-1
Query Match 100.0%; Score 2124; DB 37; Length 3117;
Best Local Similarity 100.0%; Pred. No. 1.6e-292;
Matches 2124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTGTGGCTCTACACGCGCGCCCGACGCGCGCTCTATCTGCTCATCATGCGCG 60
Db 15 ATGGCTGTGGCTCTACACGCGCGCCCGACGCGCGCTCTATCTGCTCATCATGCGCG 74
QY 61 TCGTTCGGTGTCAAGAGAACCGCGAACAGCTGCTGCTGAGATTGCTGCTGCTCGCA 120
Db 75 TCGTTCGGTGTCAAGAGAACCGCGAACAGCTGCTGCTGAGATTGCTGCTGCTCGCA 134
QY 121 CGCAAGTCCACCTCGCGCTCGGCTGTACTGCTGCGCTGCTGCTGCTGCTGCTGCTG 180
Db 135 CGCAAGTCCACCTCGCGCTCGGCTGTACTGCTGCGCTGCTGCTGCTGCTGCTGCTG 194
QY 181 ATGCTGATGCTGCTGCTGAGTTCGCGCTTGTGCTCAAGACGCGCGCTGCTGCTGCTG 240
Db 195 ATGCTGATGCTGCTGCTGAGTTCGCGCTTGTGCTCAAGACGCGCGCTGCTGCTGCTG 254
QY 241 ACTGCTGGCTGCTGCTGCTGAGTTCGCGCTGCTGCTCAAGCGCGCGCTGCTGCTGCTG 300
Db 255 ACTGCTGGCTGCTGCTGCTGAGTTCGCGCTGCTGCTCAAGCGCGCGCTGCTGCTGCTG 314
QY 301 CGCTAGCAGCAGTACGCTGAGCTGCGCTGCGACACCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 315 CGCTAGCAGCAGTACGCTGAGCTGCGCTGCGACACCTGCTGCTGCTGCTGCTGCTGCTG 374
QY 361 AAGTCCGCTACCTCCATCAAGAGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 375 AAGTCCGCTACCTCCATCAAGAGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 434
QY 421 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 435 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 494
QY 481 GCTGACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 495 GCTGACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 554
QY 541 GCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 555 GCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 614
QY 601 TCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 615 TCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 674
QY 661 GCGCAAGTGGTGGTATCCACACATCGCTTCCAGGCGCGCTGCTGCTGCTGCTGCTGCTG 720
Db 675 GCGCAAGTGGTGGTATCCACACATCGCTTCCAGGCGCGCTGCTGCTGCTGCTGCTGCTG 734
QY 721 TTCAGGACACGAAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 735 TTCAGGACACGAAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 794
QY 781 GCGCAAGTGGTGGTATCCACACATCGCTTCCAGGCGCGCTGCTGCTGCTGCTGCTGCTG 840
Db 795 GCGCAAGTGGTGGTATCCACACATCGCTTCCAGGCGCGCTGCTGCTGCTGCTGCTGCTG 854
QY 841 AAGACCTACAAGAGATCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 855 AAGACCTACAAGAGATCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 914
QY 901 ACTGCTGCGCCCAACTACGCGACCGGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 915 ACTGCTGCGCCCAACTACGCGACCGGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974


```
Db 481 TTCACCAAGCCCAAGTCGGTGGCTATCCACAACATCGCTTCCAGGGCCGCGATGG 540
QY 712 GAGGAGGCTTTCAAGGACAGAAGCTGCCCGAGCGCCCTTTGACAAAGCTGGCTTCTCG 771
Db 541 GAGGAGGCTTTCAAGGACAGAAGCTGCCCGAGCGCCCTTTGACAAAGCTGGCTTCTCG 600
QY 772 GACGGCTATGCGAAGGTTTACACTGAGGCGCACCCCGATGAGGAGGACGAGAGCCCGCG 831
Db 601 GACGGCTATGCGAAGGTTTACACTGAGGCGCACCCCGATGAGGAGGACGAGAGCCCGCG 660
QY 832 CTGACGGGAAAGCACTACAAGAAGATCAACTGGCTGAAGGTGCAATTATCGCGCGCGAC 891
Db 661 CTGACGGGAAAGCACTACAAGAAGATCAACTGGCTGAAGGTGCAATTATCGCGCGCGAC 720
QY 892 AAGCTGTGATGTGTGCGCCCAACTACGCGACCGAGATCGCTGCGCATGCGCGCGGGGT 951
Db 721 AAGCTGTGATGTGTGCGCCCAACTACGCGACCGAGATCGCTGCGCATGCGCGCGGGT 780
QY 952 GTGGAGCTGGACACCGCTATCCGGCGCAAGGGCAATTGAGGGCAATTGTAAGGGATGGAC 1011
Db 781 GTGGAGCTGGACACCGCTATCCGCGCAAGGGCAATTGAGGGCAATTGTAAGGGATGGAC 840
QY 1012 ATTGAGGAGTGGAAACCCCAAGACGACAAGTTCTGTGCGCCCTACGACCAAGACAGC 1071
Db 841 ATTGAGGAGTGGAAACCCCAAGACGACAAGTTCTGTGCGCCCTACGACCAAGACAGC 900
QY 1072 GTCTACGCGGCAAGGCGCGCGCAAGGAGGSCCTGCAGCGCGAGCTGGGCGCTGCCCTGTG 1131
Db 901 GTCTACGCGGCAAGGCGCGCGCAAGGAGGSCCTGCAGCGCGAGCTGGGCGCTGCCCTGTG 960
QY 1132 GACCCACGCGCCCGCTGTTCGCTTTCATCGCGCGCTTGAGGAGGAGAGAGGTTGGAC 1191
Db 961 GACCCACGCGCCCGCTGTTCGCTTTCATCGCGCGCTTGAGGAGGAGAGAGGTTGGAC 1020
QY 1192 ATCATCTGCGCGCGCTGCCCAAGATCGTGGCCACCCCAAGTGCAGATCGCATCTGT 1251
Db 1021 ATCATCTGCGCGCGCTGCCCAAGATCGTGGCCACCCCAAGTGCAGATCGCATCTGT 1080
QY 1252 GGTACCGGCAAGGCGCGCTAGGAGAAGCTGGTGAACGCCAATPCGACCAAGTACAGGGC 1311
Db 1081 GGTACCGGCAAGGCGCGCTAGGAGAAGCTGGTGAACGCCAATPCGACCAAGTACAGGGC 1140
QY 1312 CGCGCCAAAGCGGTGTCAGTTCTCGCGCGCTTGGCGCACATGCTCACCGCGCGGCC 1371
Db 1141 CGCGCCAAAGCGGTGTCAGTTCTCGCGCGCTTGGCGCACATGCTCACCGCGCGGCC 1200
QY 1372 GACTTCATGTGTGTCGCTCGCGCTTGGAGCGCTTGGCGCTGATCCAGCTGCAGCGCATG 1431
Db 1201 GACTTCATGTGTGTCGCTCGCGCTTGGAGCGCTTGGCGCTGATCCAGCTGCAGCGCATG 1260
QY 1432 CACTACCGTACCGTGGCGGTGGTAGCTTCCACCGGGCGCTGGTCGACACCGTCAAGGAG 1491
Db 1261 CACTACCGTACCGTGGCGGTGGTAGCTTCCACCGGGCGCTTGGTCGACACCGTCAAGGAG 1320
QY 1492 GGGCTACCGGCTTCCACATGGCGCCCTGAACCCCGACAAAGCTGGAGGAGGCTGACGCC 1551
Db 1321 GGGCTACCGGCTTCCACATGGCGCCCTGAACCCCGACAAAGCTGGAGGAGGCTGACGCC 1380
QY 1552 GACGCCCTGGCGCGCACCGTGGCGCTGCGCAGAGGTGTTTGGGGGGCGCGCTACGCC 1611
Db 1381 GACGCCCTGGCGCGCACCGTGGCGCTGCGCAGAGGTGTTTGGGGGGCGCGCTACGCC 1440
QY 1612 GAGATGTTGGCCAACTGTCATCAGCCAGGACCTGTCTGGTCCAAAGCGCGCCGAGAGTGG 1671
Db 1441 GAGATGTTGGCCAACTGTCATCAGCCAGGACCTGTCTGGTCCAAAGCGCGCCGAGAGTGG 1500
QY 1672 GAGGCGCTGCTGGAGGAGTGGTGTACGGCAAGGCGCGCTGGCCACCGCCCAAGAGGAG 1731
Db 1501 GAGGCGCTGCTGGAGGAGTGGTGTACGGCAAGGCGCGCTGGCCACCGCCCAAGAGGAG 1560
QY 1732 GAGATCAAGTGTCCGTCGCGAGAAAGATCCCGGCGGACCTGCGCGCGCGTGTCTACGCC 1791
Db 1561 GAGATCAAGTGTCCGTCGCGAGAAAGATCCCGGCGGACCTGCGCGCGCGTGTCTACGCC 1620
```

```
QY 1792 CCCAACACCCCTGAAGCCCGTGTCCGCTTCCGTGAGGGCAAGCGCGCGCGCCCAAG 1851
Db 1621 CCCAACACCCCTGAAGCCCGTGTCCGCTTCCGTGAGGGCAAGCGCGCGCGCCCAAG 1680
QY 1852 GTCGGCACCAACCGCCCGCGCATGGGCGGTGGCGGGGACACACCCCTCGGGCCCTCG 1911
Db 1681 GTCGGCACCAACCGCCCGCGCATGGGCGGTGGCGGGGACACACCCCTCGGGCCCTCG 1740
QY 1912 CCGCGCGCGCACCCGCCAAGGTGACCACTTACAAAGCCCGCTTCCGCGCGCACCGCCCAAG 1971
Db 1741 CCGCGCGCGCACCCGCCAAGGTGACCACTTACAAAGCCCGCTTCCGCGCGCACCGCCCAAG 1800
QY 1972 CCCAAGACCGCTGCGCTCAAGCTGGCGGTGAGGCGCTCCACCACTCGACTCGGAGAAC 2031
Db 1801 CCCAAGACCGCTGCGCTCAAGCTGGCGGTGAGGCGCTCCACCACTCGACTCGGAGAAC 1860
QY 2032 GCGCGTGCCTCAAGCGCAAGCGCAAGCGTGGCTTGGCTTCCAAAGCTCGGCTGCCAAG 2091
Db 1861 GCGCGTGCCTCAAGCGCAAGCGCAAGCGTGGCTTGGCTTCCAAAGCTCGGCTGCCAAG 1920
QY 2092 CCGCTGGTCTCGCGCGCGCACCGCAAGTCCGCC 2124
Db 1921 CCGCTGGTCTCGCGCGCGCACCGCAAGTCCGCC 1953
```

RESULT 4

US-09-980-771A-8

; Sequence 8, Application US/09980771A

; GENERAL INFORMATION:

; APPLICANT: CNRS

; TITLE OF INVENTION: STARCH GRANULES CONTAINING A RECOMBINANT POLYPEPTIDE OF INTEREST

; FILE REFERENCE: METHOD OF OBTAINING THEM AND THEIR USES

; CURRENT APPLICATION NUMBER: US/09/980,771A

; CURRENT FILING DATE: 2001-11-15

; PRIOR APPLICATION NUMBER: FR 99/06494

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

; LENGTH: 1593

; TYPE: DNA

; ORGANISM: artificial sequence

; FEATURE:

; OTHER INFORMATION: fragment of the complete cDNA coding for the GBSSI of Chlamydomonas reinhardtii

; NAME/KEY: CDS

; LOCATION: (1)..(1593)

; OTHER INFORMATION:

US-09-980-771A-8

Query Match

Best Local Similarity 75.0%; Score 1593; DB 37; Length 1593;

Matches 1593; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 172 GCGCTGGACATCGTGATGGTTGCTGCTGAGGTGGCCCTTGGTCCAAAGCGGGGGCGCTG 231
Db 1 GCGCTGGACATCGTGATGGTTGCTGCTGAGGTGGCCCTTGGTCCAAAGCGGGGGCGCTG 60
QY 232 GCGCATGTGACTGTGGGCTGCTATTGAGCTGGTCAAGCGCGGCCCGGCTCATGACC 291
Db 61 GCGCATGTGACTGTGGGCTGCTATTGAGCTGGTCAAGCGCGGCCCGGCTCATGACC 120
QY 292 ATTGCCCTCGCTACGACCACTACGCTGACGCCCTGGGACACCTCGTGGTGGTGACATC 351
Db 121 ATTGCCCTCGCTACGACCACTACGCTGACGCCCTGGGACACCTCGTGGTGGTGACATC 180
QY 352 ATGGGCGAGAAGTCCGCTACTTCCATCAAGAAAGGCGGTGCACCGGCTGTGGATT 411
Db 181 ATGGGCGAGAAGTCCGCTACTTCCATCCATCAAGAAAGGCGGTGCACCGGCTGTGGATT 240
QY 412 GACCACCCCTGGTCTCTGGCCCAAGGTCTGGGGCAAGACCGGCTCCAAAGCTGTACGGCCCC 471
```

Db 241 GACACCCCTGGTTCTTGGCCAAAGGCTGGGCAAGACCGGCTCCAAAGCTGTACGGCCC 300
QY 472 CGTCCCGCGCTGACTACCTGGGCAACCAAGCGCTTGCGCCCTGTCTGCAAGCGCGCT 531
Db 301 CGTCCCGCGCTGACTACCTGGGCAACCAAGCGCTTGCGCCCTGTCTGCAAGCGCGCT 360
QY 532 ATTGAGGCTGCCCGCGTGTGCTTCCGCTTCCGCGCCGCGGCGAGGACTGCTCTGCTGGCCCAAC 591
Db 361 ATTGAGGCTGCCCGCGTGTGCTTCCGCTTCCGCGCCGCGGCGAGGACTGCTCTGCTGGCCCAAC 420
QY 592 GACTGGCACTCCGCGCTGTGGCGCTTCTGCTGAAGGACGAGTACAGCCCAAGGGCCAG 651
Db 421 GACTGGCACTCCGCGCTGTGGCGCTTCTGCTGAAGGACGAGTACAGCCCAAGGGCCAG 480
QY 652 TTCACCAAGGCAAGTGGCTGTGCTATCCACAACATCGCTTCCAGGGCGCGCATGTGG 711
Db 481 TTCACCAAGGCAAGTGGCTGTGCTATCCACAACATCGCTTCCAGGGCGCGCATGTGG 540
QY 712 GAGGAGGCTTTCAGGACACGAGCTGCCCGCGCTTTCAGAGCTGGGCTTCTCG 771
Db 541 GAGGAGGCTTTCAGGACACGAGCTGCCCGCGCTTTCAGAGCTGGGCTTCTCG 600
QY 772 GAGGCTATGCCAAGGTTTACATGAGGCGCACCCCATGATGAGGAGGACGAGAACCCCG 831
Db 601 GAGGCTATGCCAAGGTTTACATGAGGCGCACCCCATGATGAGGAGGACGAGAACCCCG 660
QY 832 CTGACGGGAAGACCTTCAAGAGATCAACTGGCTGAAGGTGCAATTATCGCCGCGCAC 891
Db 661 CTGACGGGAAGACCTTCAAGAGATCAACTGGCTGAAGGTGCAATTATCGCCGCGCAC 720
QY 892 AAGTGTGTAGTGTGCGCCAACTACGCGACGAGATCGCTGCCGATGCGCGCGGCT 951
Db 721 AAGTGTGTAGTGTGCGCCAACTACGCGACGAGATCGCTGCCGATGCGCGCGGCT 780
QY 952 GTGAGCTGACACCGCTATCCGCGCAAGGCAATTGAGGCAATTGAGGCAATTGAGGCAATTGAG 1011
Db 781 GTGAGCTGACACCGCTATCCGCGCAAGGCAATTGAGGCAATTGAGGCAATTGAGGCAATTGAG 840
QY 1012 ATTGAGAGTGGAAACCCCAAGACCGACAAGTTCCTGTCTGCGCCCTACGACCAAGACAGC 1071
Db 841 ATTGAGAGTGGAAACCCCAAGACCGACAAGTTCCTGTCTGCGCCCTACGACCAAGACAGC 900
QY 1072 GTCTACCGCGCAAGGCGCGCGCAAGGAGGCGCTGACGCGAGCTGGCGCTGCTGTG 1131
Db 901 GTCTACCGCGCAAGGCGCGCGCAAGGAGGCGCTGACGCGAGCTGGCGCTGCTGTG 960
QY 1132 GACCCACCGCGCGCTGTGCTTCTATCGCGCGCTTGGAGGAGCAAGAGGTGTGAGC 1191
Db 961 GACCCACCGCGCGCTGTGCTTCTATCGCGCGCTTGGAGGAGCAAGAGGTGTGAGC 1020
QY 1192 ATCATCTGCGCGCGCTGCCCAAGATCCTTGGCCACCCCAAGGTGCAAGTTCGCAATCCTG 1251
Db 1021 ATCATCTGCGCGCGCTGCCCAAGATCCTTGGCCACCCCAAGGTGCAAGTTCGCAATCCTG 1080
QY 1252 GSTTACCGCAAGGCGCGCTTACGAGAGCTGTGTAAGCCATCGGCCCAAGTACAGGGC 1311
Db 1081 GSTTACCGCAAGGCGCGCTTACGAGAGCTGTGTAAGCCATCGGCCCAAGTACAGGGC 1140
QY 1312 CGGCGCAAGGCGCTGTCAAGTTCCTGCGCGCGCTTGGCGCACATGCTTACCGCGCGCGC 1371
Db 1141 CGGCGCAAGGCGCTGTCAAGTTCCTGCGCGCGCTTGGCGCACATGCTTACCGCGCGCGC 1200
QY 1372 GACTTATGCTGTGCGCTTGGAGCCCTTGGAGCCCTGTGAGCTGTGACGCTGACGCGCATG 1431
Db 1201 GACTTATGCTGTGCGCTTGGAGCCCTTGGAGCCCTGTGAGCTGTGACGCTGACGCGCATG 1260
QY 1432 CACTAGGTAACGCTGCGGTGTAGCTTCCACCGCGCGCTTGTGCGACCCGCTCAAGGAG 1491
Db 1261 CACTAGGTAACGCTGCGGTGTAGCTTCCACCGCGCGCTTGTGCGACCCGCTCAAGGAG 1320
QY 1492 GCGCTACCGGCTTTCACATGAGGCGCGCTTGAACCCCGACAAAGCTGGACGAGGCTGACGCC 1551

Db 1321 GCGCTACACGGCTTTCACATGGGCGCGCTGAAACCCCGACAAAGCTGGAGAGGCTGACGCC 1380
QY 1552 GACGCCCTGGCGCGCACCGTGGCGTCCAGCGAGGTGTTTGGGGGGCGCGCTACGCC 1611
Db 1381 GACGCCCTGGCGCGCACCGTGGCGTCCAGCGAGGTGTTTGGGGGGCGCGCTACGCC 1440
QY 1612 GAGATGTGGCCAACTGCATCAGCAGGACCTGTCTGTGTCACAAAGCCGCCAGAGTGG 1671
Db 1441 GAGATGTGGCCAACTGCATCAGCAGGACCTGTCTGTGTCACAAAGCCGCCAGAGTGG 1500
QY 1672 GAGGCGCTGTGGAGGAGTGTGTACGCGAAGGCGCGCTGGCCACCGCCCAAGAGAG 1731
Db 1501 GAGGCGCTGTGGAGGAGTGTGTACGCGAAGGCGCGCTGGCCACCGCCCAAGAGAG 1560
QY 1732 GAGATCAAGTGGCCGCTTGGCGAAGAGATGCC 1764
Db 1561 GAGATCAAGTGGCCGCTTGGCGAAGAGATGCC 1593

RESULT 5

US-09-980-771A-6

; Sequence 6, Application US/09980771A

; GENERAL INFORMATION:

; APPLICANT: CNRS

; TITLE OF INVENTION: STARCH GRANULES CONTAINING A RECOMBINANT POLYPEPTIDE OF INTER

; FILE OF INVENTION: METHOD OF OBTAINING THEM AND THEIR USES

; FILE REFERENCE: WO 99 AB CNR AMYL

; CURRENT APPLICATION NUMBER: US/09/980,771A

; CURRENT FILING DATE: 2001-11-15

; PRIOR APPLICATION NUMBER: FR 99/06494

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 1314

; TYPE: DNA

; ORGANISM: artificial sequence

; FEATURE:

; OTHER INFORMATION: fragment of the complete cDNA coding for the GBSSI of Chlamydo

; OTHER INFORMATION: as reinhardtii

; NAME/KEY: CDS

; LOCATION: (1)..(1314)

; OTHER INFORMATION:

US-09-980-771A-6

Query Match

61.9%; Score 1314; DB 37; Length 1314;

Best Local Similarity 100.0%; Pred. No. 1.9e-177;

Matches 1314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 GCGTGGACATCGTGTGCTGCTGAGTCCGCCCTTGGTCCAAAGACGGCGGCGCTG 60
Db 1 GCGTGGACATCGTGTGCTGCTGAGTCCGCCCTTGGTCCAAAGACGGCGGCGCTG 231
QY 232 GCGCATGTGACTGTGGCGCTGCCCTATTGAGCTGTGTCAGCGCGGCCACCGCTATGACC 291
Db 61 GCGCATGTGACTGTGGCGCTGCCCTATTGAGCTGTGTCAGCGCGGCCACCGCTATGACC 120
QY 292 ATTGCCCTCGCTACGACCACTAGCTGAGCGCTGGGACACCTCGGTGGTGTGTCGACATC 351
Db 121 ATTGCCCTCGCTACGACCACTAGCTGAGCGCTGGGACACCTCGGTGGTGTGTCGACATC 180
QY 352 ATGGCGGAGAGGTCCTGCTACTTCCACTCCATCAAGAGGCGGTGACCGGCTGTGAT 411
Db 181 ATGGCGGAGAGGTCCTGCTACTTCCACTCCATCAAGAGGCGGTGACCGGCTGTGAT 240
QY 412 GACACCCCTGGTTCCTGGCCCAAGTCTGGGCAAGACCGCTCCAGCTGTACGGCC 471
Db 241 GACACCCCTGGTTCCTGGCCCAAGTCTGGGCAAGACCGCTCCAGCTGTACGGCC 300
QY 472 CGTCCGGCGCTGACTACCTGGACCAACCAAGCGCTTGGCCCTGTGTTCTGCAAGCGGCT 531
Db 301 CGTCCGGCGCTGACTACCTGGACCAACCAAGCGCTTGGCCCTGTGTTCTGCAAGCGGCT 360

Db 687 CGGACCTACGGGGAAGACGCTGGTGTCTGTGTGCAAGACTGGGACACGGGCTCTGGC 746
Qy 615 CGTCTCTGCTGAAGGACGAGTACACAGCCCAAGGCGCAGTTTACCAAGGCCCAAGTGGTGTCT 674
Db 747 CTGCTACCTCAAGAGCAACTACACAGTCCAGTGGCACTATAGGACGCCCAAGGTAGGCTT 806
Qy 675 GGTATATCCCAACATCGCTTCCAGGCGCCGATGTGGAGGAGGCTTCAAGGACACGAA 734
Db 807 CTGATCCCAACATCTCGTATCAGGCGCCGCTTCTCTTCGACGACTTCCGCGAGCTCAA 866
Qy 735 GCTGGCCCGAGCGCCCTTTGACAAGCTGGCTCTCGGACGGCTATGCCAAGGTTTACAC 794
Db 867 CTTGCCCGAGAGTTCAAGTGTCTTCGACTTTCATCGACGGGTACGACAA 918
Qy 795 TGAGGCCACCCCATGAGGAGGAGGACGAGAGGCCCGCTGACGGGAAAGACCTACAAGAA 854
Db 919 -----CCGTGGAGGGGCGC-----AA 935
Qy 855 GATCAACTGCTGAAGGTTGGCATATTCGCGCGCGACAGCTGGTACTGTGTGGCCCAA 914
Db 936 GATCAACTGGATGAAGGCCCGGATCTCTGAGCGGACAAAGTGTCTACGGTGAAGCCCTA 995
Qy 915 CTACGCGACGACGATCGCTCCGATCGCGCGGCTGTGAGCTGGACACCTCATCCG 974
Db 996 CTACGCGGAGGACTCATCTCCGCGAAGCCAGGGGTGCGAGCTGACAACTCATGCG 1055
Qy 975 CGCAAGGGCAITGAGGGCAATTGTGAACGSCATGAGATGAGAGTGAACCCCAAGAC 1034
Db 1056 CCTCAGGGGATCACCAGGATGCTCAACGSCATGGAGTCAAGGAGTGGGACCCGCCAA 1115
Qy 1035 CGACAGTTCTCTGCGCCCTACGACCAAGACAGCTGTACCGCGGCAAGCGCGCGC 1094
Db 1116 GGCAAGTTCTCTACCGCCAACTTACGACGCTCAACACCGGTTGAGGGGAAGCGCTGAA 1175
Qy 1095 CRAAGAGGCGCTGCAGGCGGAGCTGGCGCTGTGTGGACCCACCGCCCGCTGTTCG 1154
Db 1176 CRAAGAGGCGCTGCAGGCGGAGTGGGGTGTGCGGTGGACCGAAGTGGCCCTGTGGC 1235
Qy 1155 CTTCTATCGGCGCTGGAGAGCAGAGGAGTGTGACATCATCTGGCGCCCTGCCCAA 1214
Db 1236 CTTCTATCGGCGCTGGAGAGCAGAGGAGTGTGACATCATCTGGCGCCCGGATGATG 1295
Qy 1215 GATCTGTGCGCCACCCCAAGTGCAGATCGCCATCTTGGTACCGGCAAGCGCCCTACGA 1274
Db 1296 GATCTTGAAGAGGAGGAGCTCCAGATCGTTCCTTGGGCAACCGGGAAGAAGTTGA 1355
Qy 1275 GAAGCTGTGAAGCGCATCGGACCAAGTACAAGGCGCGCCCAAGGCGTGTCAAGTT 1334
Db 1356 GCGGCTGCTCAAGAGCGTGGAGGAGAGTTCCCGACCAAGTCAAGGCGCGTGTCAAGTT 1415
Qy 1335 CTCGGGCGCTGGGCGACATGCTACCGCGGCGCGGAGCTTCAATGCTGTGCGCCCTCGG 1394
Db 1416 CAACGCGCGCTGGTCTACAGATGATGGCGCGCGGACGCTGCTGCGCGTCAACAGCG 1475
Qy 1395 CTTGAGCGCTGCGGCTGATCCAGCTGCACCGCATGCTAGGTACCGTCCCGTGGT 1454
Db 1476 CTTGAGCGCTGCGGCGCTATCCAGTCCAGGGATGCGCTACGGAACGCGGTGGCGCTG 1535
Qy 1455 AGCTTCCACCGGCGCTGGTGCACACCGTCAAGGAGGCGCTACCGGCTTCCACATGGG 1514
Db 1536 CGCGTCCACCGGCGCTGCTCGACACGATCATGGAAGCAAGACCGGGTTCACATGG 1595
Qy 1515 CGCCCTGA-----ACCCGCAAGCTGACAGGCTGACCGGAGCGCCCTGCGCGCAC 1568
Db 1596 CACCTTCAGCTGCAAGCTGGTGGAGCGCGCGGACGTGAAGAAGGTGGTGACAC 1655
Qy 1569 CTTGGCGCTGCGAGGAGTGTTCGCGGCGCGCGCTACCCCGAGATGGTGGCCAACTG 1628
Db 1656 CTTGAAGCGCGCTCAAGGTCGTGGACGCGCCAGCTTACCATGATGATGTTCAAGAACTG 1715
Qy 1629 CATACGCGAGGACCTGCTCTGTGTCACGCGCCGCAAGAGTGGGAGGCGCTGCTGGAGGA 1688
Db 1716 CATGATCCAGGATCTCTCTCTGGAAGGGGCGCCAGCAAGAACTGGGAGGAGCTCTTCTGGA 1775

Qy 1689 GGTGTGTACGGCAAGGCGCGCTGGCCACCGCCCAAGAGGAGGATCAAGGTGCCGT 1748
Db 1776 ACTGGG---GTCAGGGGAGGAGCCAGGGGTATCGGCGAGAGATTCGCCGCTCGC 1832
Qy 1749 TGGCGAGAAGATCCCG 1765
Db 1833 CATGGAGAAGTCGCG 1849

RESULT 7
US-07-771-351-5
; Sequence 5, Application US/0771351
; GENERAL INFORMATION:
; APPLICANT: LILLIS, MARCELLA
; APPLICANT: CHEN, CHE-HONG
; APPLICANT: GROSS, PAUL R.
; APPLICANT: FANG, LAWRENCE Y.
; TITLE OF INVENTION: GENETIC SEQUENCES FOR ALTERING
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
; STREET: 5 PALO ALTO SQUARE, SUITE 400
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/771,351
; FILING DATE: 19911004
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY, RICHARD L.
; REGISTRATION NUMBER: 30092
; REFERENCE/DOCKET NUMBER: LFCP-015/0005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-494-7622
; TELEFAX: 415-857-0663
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2201 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; STRAIN: WHEAT VAR. PROBRAND755
; IMMEDIATE SOURCE:
; CLONE: PSOG190
; US-07-771-351-5

Query Match 24.8%; Score 527.4; DB 3; Length 2201;
Best Local Similarity 61.3%; Pred. No. 9.4e-66;
Matches 1008; Conservative 0; Mismatches 541; Indels 96; Gaps 6;
Qy 163 GCACHTTGGCGCTGGACATCGTGTGCTGCTGAGTGGTGTGCTGAGTGGCCCTTGGTCCAAAGACG 222
Db 340 GGCAGCGGGGCGCATGAACCTCGTGTTCGTGCGCGGAGATGGCGCCCTGGAGCAAGACC 399
Qy 223 GCGCGCTCGGGCATGACTGTGGCTGCCTTATTGAGTGTGCTCAAGCGCGCCACCGC 282
Db 400 GCGGGCTCGGAGACGTCTCGGGGGCTTCCGCCAGCCATGCGCCCAACGCCACCGG 459
Qy 283 GTCATGACCATTTGCCCTCGCTACGACCACTAGCTGACGCTTGGGACACTCGGTGGTGC 342
Db 460 GTCATGTCATCTCCCGCGCTACGACCACTAGCAAGGACGCTTGGGACACCAACCGCTGTC 519

US-07-771-351-3

Query Match	24.7%;	Score 523.8;	DB 3;	Length 2279;
Best Local Similarity	61.2%;	Pred. No. 3e-65;		
Matches 986;	Conservative 0;	Mismatches 532;	Indels 93;	Gaps 5;
QY	163	GCACCTTGGCGCTGGACATCGTATGGTTGCTGCTGAGGTGCGCCCTTGCTTCCAAAGACG	222	
DB	391	GGCAGCGGCGCATGACCTCGTTCGTTCGCGCGGAGATGGCGCCCTTGAGCAAGACT	450	
QY	223	GGCGGCTTGGCGATGTAAGTGGTGGCTGCTATTGAGCTGGTCAAGCGGCGCACCGC	282	
DB	451	GGCGGCTTGGCGAGCTCCTCGGGGCGCTCCCGCGCATGGCGGCCAACGGTCAACGG	510	
QY	283	GTCAATGACCATTCGCTTCGTACAGCACCTGACGTACGCTGACGCTGGGACACCTCGTGGT	342	
DB	511	GTCAATGCTCATCTCCCGCGCTACGACCATGTAAGGAGCGGCTGGGACACCGAGTCA	570	
QY	343	GTGACATCATGG-----GGAGAAGTTCGCTACTTCCACTTCCATCAAGAAG	390	
DB	571	TCCGAGATCAAGGTTCGTGACAGTACGAGAGGTGAGGTACTTCCACTGCTACAAGCGC	630	
QY	391	GGCGTGACCGCGTGTGATGTGACCAACCTGGTTCCTGGCCAGGTCTGGGGCAAGACC	450	
DB	631	GGGTGACCGCGTTCGTGCAACACCGCTGCTTCTGGAGAAGTTCGGGGGCAAGACC	690	
QY	451	GGCTCCAAGTGTACGGCCCGCGCTCGGGGCGTCACTACCTGGACAACCAACAGCGTTC	510	
DB	691	AAGGAGAAGATCATGGAACCGGCGCGGCAACCGACTACGAGGACACCAAGCAGCTTC	750	
QY	511	GCCTGTTCGCAAGGCGCGTATTGAGGCTGCGCGCGTGTGCGCCCTTCGGGCC-----	563	
DB	751	AGCCTTCTCTGCCAGGCAAGCATGAGTGGTCCAGGATCCTCGACCTCAACAACAACCA	810	
QY	564	-----CGCGGAGGACTGGTCTTCGTGCGCAACGACTGCACTCCGCGC-----	606	
DB	811	CACCTTCTCGACCTACGGGAGAGCGTGGTTCGTGTCACAGACTGCAACAGGCGC	870	
QY	607	CTGGTGCCGCTGCTGCTGAAGGAGGAGTACACGCCCAAGGCCAGTTCACCAAGGCCAAG	666	
DB	871	CTTCTGCCCTGCTACCTCAAGAGCAACTACGAGTCCAATGSCATCTATAGGACGCGCAAG	930	
QY	667	TCGGTGTGCTGCTATCCACACATCGCTTCCAGGCGCGATGTGGGAGGAGCTTCAAG	726	
DB	931	GTGGCATCTGCAATCCAAACATCTGATACAGGCGCGCTTCTCTTCGACGACTTCGCG	990	
QY	727	GACACGAAGCTGCCCGGCGGCTTTGACAAGCTGGCCTTCTCGGACGGCTATGCCAAG	786	
DB	991	CAGCTCAACTGCGCGACAGGTTCAAGTCGCTTCGACTTCATCGACGGCTAGACAAG	1050	
QY	787	GTTTACACTGAGGCCACCCCATGGAGGAGAGAGAAGCCCGCTGACGGGGAAGACC	846	
DB	1051	-----CGGTGGAGGCGCGC-----	1065	
QY	847	TACAAGAAGATCAACTGGCTGAAGGTTGGCATTATCGCGCGACAAAGCTGGTGA	906	
DB	1066	-----AGATCACTGGATGAAGCGCGGATCTTCGAGCGCGACAGGTGTGCACTGTG	1119	
QY	907	TCGCGCAACTACGCGACCGAGATCGCTGCGATGCGCGCGCGCTGTGAGCTGGACACC	966	
DB	1120	AGCCCCCTACTATGCTGAGGAGCTAAATCTCTGGCGAAGCCAGGGGCTCGGAGCTCAAC	1179	
QY	967	GTCAATCGGCGCAAGGCGATTGAGGCGATTGTGAACGGCATGACATGAGGAGTGAAC	1026	
DB	1180	ATCATGGCGCTCACTGGGATCAACCGCATCGTCAACGGCATGACGCTCAACGAGTGGGAC	1239	
QY	1027	CCCAAGACCGACAAGTTCCTGCTGCGCGCTACGACCAAGAAACAGCGTCTACGCGGCAAG	1086	
DB	1240	CCCATCAAGGACAAGTTCCTCACCGTCACTACGAGTCAACACCGCGGTGAGGGGAAG	1299	
QY	1087	GGCGCGCAAGAGGCGCTGCAAGCGGAGCTGGGCTGCTGTGAGACCCCAACCGGCCCC	1146	
DB	1300	GGCTCAACAGAGGCGCTGCAAGCGGAGGTGGGCTGCGGCTGACCGCAAGGTGCC	1359	

```

RESULT 9
US-09-731-166-3
; Sequence 3, Application US/09731166
; GENERAL INFORMATION:
; APPLICANT: Sewalt, Vincent J. H.
; APPLICANT: Singletary, George W.
; TITLE OF INVENTION: Production of Modified Polysaccarides
; FILE REFERENCE: 35718/206348
; CURRENT APPLICATION NUMBER: US/09/731,166
; CURRENT FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: 60/169,993
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: UDP-glucose starch glycosyl transferase
; OTHER INFORMATION: Accession No. 22509
; NAME/KEY: CDS
; LOCATION: (1)...(1818)
US-09-731-166-3

```

Query Match 22.5%; Score 478; DB 29; Length 1818;
Best Local Similarity 59.8%; Pred. NO. 1e-58;
Matches 986; Conservative 0; Mismatches 565; Indels 99

Db 2201 CCTGGGCTCGCCGCGCAGCGCGCGGGGATCG 2232
||| || | ||| || | ||

Search completed: June 7, 2003, 01:55:08
Job time : 3358.08 secs

Db 518 CCGAGGAGAAGATCTACGGGCGCTGTGCTGGAACGAGCTACAGGACAAACAGCTGCGGT 577
QY 509 TCGCCCTCTTCGAAGCGCGCTATTGAGGCTGCGCGGTGCTGCCCTTCGCG----- 561
Db 578 TCAGCCTGTATGCCAGGACAGCTTGAGGCTCCAAAGGATCTGAGCCTCAACAACAACC 637
QY 562 -----CCGCGGAGGACTGGGTCTTCGTGGCCCAACGACTGCGCATCCG 604
Db 638 CATACCTCTCCGAGACATACGGGAGGAGCTGTGCTGCTGCAACGACTGGCACCG 697
QY 605 CCTGTGTCCTGCTGCTGAGGAGGAGTACAGCCCAAGGCCAGTTTACCAAGGCCA 664
Db 698 GCGCTCTCTGCTGCTACCTCAAGAGCAACTACAGTCCACGCGCATCTACAGGAGCA 757
QY 665 AGTGGTGTGCTATFCCCAACATCGCTTCCAGGGCGCATGTGGGAGGAGGCTTCA 724
Db 758 AGACCGCTTCTGATCCACACATCTCTACAGGGCGGTTCGCTTCTCGACTACG 817
QY 725 AGGACACAAAGTGGCCCAAGGCGGCTTTGACAAGGTGGCTTCTCGGACGGCTATGCCA 784
Db 818 CGGAGCTGAACCTCCCGAGAGATTCAAGTCTGCTTCTGATTTTCATCGACGGCTACGAGA 877
QY 785 AGGTTTACACTGAGGCCACCCCATGAGGAGGACGAGAGCCCGCTGACGGGAAAGA 844
Db 878 AG-----CCGTGGAAGG----- 890
QY 845 CCAACAAGAAGTCAACTGGCTGAAGGTGGCATTTATCGCGCGGACAAAGCTGGTACTG 904
Db 891 ----CGGAAGATCACTGATGATGAGGCGGGATCTCGAGCGGACAGGTTCTTACCG 946
QY 905 TGTGGCCCAACTACGGGACCGAGATCGCTGCGGATGCCCGCGGCGGTGTGGAGTGGACA 964
Db 947 TCAGGCCCTACTACGCGGAGGAGCTCATCTCGGGCATCGCCAGGGCTCGGAGCTCGACA 1006
QY 965 CGGTATCGCCCAAGGCGATTGAGGCGATTGTGAACGCGCATGGACATTGAGGATGGA 1024
Db 1007 ACATATGCGCTCACCGGCATCACCGGCATCGTCAACGGCATGGACGTCAGGAGTGGG 1066
QY 1025 ACCCAAGACCGACAAGTTCTCTGCTGCGCCCTACGACAGAACAGCTCTACGCGGCA 1084
Db 1067 ACCCAAGGAGGACAGTATACGCGGTGAAGTACGACGCTGTCAGCGCGGTGGAGGCCA 1126
QY 1085 AGGCGCGCCCAAGGAGGCGCTGAGGCGGAGCTGGGCTGCTGTGGACCCCAACGCGCC 1144
Db 1127 AGGCGCTGAACAAGGAGGCGCTGAGGCGGAGGTGCGGCTCGCGGTGGACCGGAAATCC 1186
QY 1145 CCGTGTTCCTTATCGCGCGCTTGGAGGAGCAGAAAGGTGTGGACATCATCTGCGCG 1204
Db 1187 CGCTGGTGGCTTATCGGCGGCTGGAAGAGCAGAAAGGCGCCCGCTACGTCATGGCGCG 1246
QY 1205 CCTGCGCCAAAGATCTCGCCACCCCGCAAG---GTGAGATCGCCATCTCTGGGTACCGCA 1261
Db 1247 CCATCCCGAGCTCATGGAGATGGTGGAGGACGTGCAGATCGTTCTGTGGGACCGGCA 1306
QY 1262 AGGCGCGCTACAGAGCTGTGAGCGCATCGGCACCAAGTACAGAGGCGCGGCCAAG 1321
Db 1307 AGAAGAAGTTCGAGGCGCATGCTCATGAGCGCGGAGGAGAGTTCCCAAGGCAAGGTGCGG 1366
QY 1322 GGTGGTCAAGTTCTCGGCGCGCTTGGGCGCATGCTACCGCGCGCGGCTTCATGC 1381
Db 1367 CGCTGGTCAAGTCAAGCGCGCTGCGCGCACCATCATATGCGCGCGCGGCTGCTCG 1426
QY 1382 TGTGCGCTTCGCTTCGAGCGCTGCGGCTGATCCAGCTGACGCGCATGCACTACGGTA 1441
Db 1427 CGGTCAACAGCGCTTCGAGCGCTGCGGCTCATCCAGCTGCGAGGGGATGCGATACGGAA 1486
QY 1442 CGGTGCGCTGTAGCTTCCACCGGCGCTTGGTGACACCGTCAAGGAGGGGTACCG 1501
Db 1487 GCGCTTGCCTCTCGCGCTCCACCGGTGGACTGTGCGACACCATCATCATCAAGGCAAGACG 1546
QY 1502 GGTTCACATGGCGCGCTCA-----ACCCGCAAGCTGGAGGAGGCTGACGCGGAG 1555
Db 1547 GGTTCACATGGCGCGCTCAGGCTGACTGTGAACGTGCTGGAGCGCGGAGCTCAAGA 1606

QY 1556 CCTGTGCGCGCACCGTGGCGGCTGCCAGCGAGGTGTTTGGGGGCGCGCTACCCGAGA 1615
Db 1607 AGGTGCCACCACTTGCAGCGCGCATCAAGTGGTGCACACCGCGGTACGAGAGA 1666
QY 1616 TGTGCGCCAACTCATCAGCAGGAGCTGTCTGTGGTCCAAAGCGCCGCAAGTGGGAGG 1675
Db 1667 TGTGTAGGAACATCATATCCAGGATCTCTCTGGAAGGCGCTTGCAGAACTGGGAGA 1726
QY 1676 GCTGTGTGAGGAGGTGGTGTACGCGAAGCGCGCGGTGCCACCGCCCAAGAGGAGAGA 1735
Db 1727 AGTGTGTGTCAAGCTCGGGGTGCGC---GGCGGCGAGCGAGGTTGAAGGCGGAGAGA 1783
QY 1736 TCAAGTGTGCGGCTGCGGAGAAATCCCG 1765
Db 1784 TCGCGCGCTCGCCAAGGAGAGCTGGCG 1813

RESULT 2

US-10-228-063-7

; Sequence 7, Application US/10228063

; GENERAL INFORMATION:

; APPLICANT: Lanahan, Mike

; TITLE OF INVENTION: Self-processing Plants and Plant Parts

; FILE REFERENCE: 109846.317

; CURRENT APPLICATION NUMBER: US/10/228,063

; CURRENT FILING DATE: 2002-12-12

; NUMBER OF SEQ ID NOS: 60

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 1818

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: synthetic

US-10-228-063-7

Query Match

Best Local Similarity 22.5%; Score 478; DB 9; Length 1818;

Matches 986; Conservative 0; Mismatches 565; Indels 99; Gaps 7;

QY 161 GTGCCACTTGGCGCTGGACATCGTATGTTGCTGCTGAGGTGCGCCCTTGTTCAGA 220
Db 218 GCGCCAGCGCGCATCAAGCTCTCTGCTGCGCGGAGATGGCGCGGTGGCAAGA 277
QY 221 CGGCGCGCTGGCGGATGTGACTGTGGCTGCTATTGAGTGGTCAAGCGGCGCAC 280
Db 278 CCGGAGGCTTCGCGGAGCTCTCGCGGCTGCGCGGCTGCGCGGCTGCGCGGCGCAC 337
QY 281 GGTGATGACATTTGCCCTCGCTACGACCACTACGCTGAGCTGGGACACTCGGTGG 340
Db 338 GTGCTATGTTGCTCTCTCCCGCTACGACCACTACAAGGACGCTGGGACACCGCTCG 397
QY 341 TGTGTGACATCA-----TGGGCGAGAGGTCCGCTACTTCCACTCCATCAAGA 388
Db 398 TGTCCGAGATCAAGATGGGAGCGGTACGAGACGCTAGGTTCTTCCACTGTAACAGC 457
QY 389 AGGCGCTGCACCGCTGTGGATTGACACCGCTTGTCTGTCGCAAGGCTCTGGGCAAGA 448
Db 458 GCGGAGTGGACCGGCTGTGTTGACACCACTGTTCTGAGAGAGGTTTGGGGAAGA 517
QY 449 CCGGCTTCAAGCTGTACGGCGCGCTCGCGGCGCTGACTACTGAGCAACCAAGCGCT 508
Db 518 CCGAGGAGAAGATCTACGGGCTGTGCTGGAACGAGCTACAGGAGCAACCAAGCTCGGT 577
QY 509 TCGCCCTGTCTGCAGGCGGCTATTGAGGCTGCGCGGCTGCTTCCGCTTCGCG----- 561
Db 578 TCAGCTGTGTATCCAGGCGGAGCTTGAAGCTCCAAGGATTCCTGAGCTTCAACAACAC 637
QY 562 -----CCGCGGAGGAGTGTGCTCTTCGTTGCGTCTGCTCAAGGACTGCAACCG 697
Db 638 CATACTTCTCCGACCATAGGGGAGGAGCTGCTGTTGCTCTGCTCAAGGACTGCAACCG 697


```
QY 845 CCTACAAGATCAACTGGCTGAAGGTGGGCAATATCCGCCCGCCGACAAAGCTGGTGACTG 904
Db 1080 ---GCCGAAGATCAACTGGATGAAGGCGGGATCCTCGAGGCGGACAGGCTCCACACG 1136
QY 905 TGTCTGCGCAACTACCGACCGAGATCGCTGCCGATGCGCGCGCGGTGTGGAGCTGGACA 964
Db 1137 TCAGCCCTTACTACCGCGAGAGCTCACTCCGGCATGCCAGGGCTCCGAGCTCGACA 1196
QY 965 CGGTCAATCGCCGCAAGGCAATGAGGGCAATGTGAACGGCATGGACATTTGAGGATGGA 1024
Db 1197 ACATCATGCGCTCACCGGCATCACCGGCATCGTCAACGGCATGGAGCTCAGCGATGGG 1256
QY 1025 ACCCAAGACCAAGATTCCTGTCTGCGCCCTAGCACCAGACAGCGTCTAGCGGGCA 1084
Db 1257 ACCCCAGAGGACCAAGTACATCGCCGTAAGTACGAGTGTGCGACGGCGGTGGAGGCA 1316
QY 1085 AGGCGCGCCGCAAGGAGGCGCTGCAGGCGAGCTGGCGCTGCTGTGGACCCACCGCC 1144
Db 1317 AGGCGCTGAACAGAGGCGCTGCAGGCGAGGTGCGGCTCCGGTGGACCGAATCC 1376
QY 1145 CCCTGTCCTTCATCGGCGCCCTGGAGGACAGAGGCTGTGGACATCATCTGCGCG 1204
Db 1377 CGCTGGTGGCTTCATCGCAGCTGGAAGACAGAGGCGCCCGACGCTCATGGCGCG 1436
QY 1205 CCTGCCCAAGATCTGGCCACCCCAAG---GTGAGATCGCCATCTCTGGGTACCGCA 1261
Db 1437 CCATCCCGAGCTCATGAGATGTTGGAGAGCTGCAGATCGTCTCTGTGGGCGAGGCA 1496
QY 1262 AGGCGCGCTTACGAGAGCTGTGAACCGCATCGGCACCAAGTACAAAGGCGCGGCAAG 1321
Db 1497 AGAAGAAGTTCAGCGCATGCTCATAGCGCGGAGGAGAGTTCACGCAAGGTGCGCG 1556
QY 1322 GCGTGGTCAAGTTCGCGGCGCCCTGGGCGACATGCTACCGCGCGCGCGCTTCAATGC 1381
Db 1557 CGGTGGTCAAGTTCAGCGCGGCGTGGGCGACCATCATATGCGCGCGCGCTGCTCG 1616
QY 1382 TGGTGGCTTCGCGCTTCGAGCGCTCGGCGCTGATCCAGCTGCAGCGCATGCTACGTA 1441
Db 1617 CGTCAACGCGCTTCAGCGCTCGGCGCTCATCCAGCTGCAGGGATGCGCATCGGAA 1676
QY 1442 CGGTGCGCGGTGATGCTTCAACCGCGCGCTGCTCGACACCGTCAAGGAGGCGTCAACG 1501
Db 1677 CGCGCTGCGCTGCGCTTCCACCGGTGGACTGTCGACACCATCATCGAAGCGACACG 1736
QY 1502 GCTTCCATGAGGCGCGCTTGAACCGG-----ACAAGCTGACGAGGCTGACGCGAG 1555
Db 1737 GGTTCACATGGGCGCGCTCAGCTGACGTGTAACGTGCTGGAGCGCGCGAGCTCAAGA 1796
QY 1556 CCCTGGCGCGCACCGTGGCGCTGCGACGAGGTGTTTTCGGGCGCGCGCTACCGCGGA 1615
Db 1797 AGGTGGCCACCATGTCAGCGCGCATCAAGTGTGCGGACCGCGGCGTACGAGGAGA 1856
QY 1616 TGGTGGCAATGCAATGACCGAGGACCTGCTCTGTGTCGAAGCGCGCCGCAAGAGTGGAG 1675
Db 1857 TGGTGAAGAACTGCATGATCCAGGATCTCTCTGGAAGGGCGCTTCCAAAGAACTGGGGA 1916
QY 1676 GCTGCTGAGGAGTGTGTACGCGAAGGGCGGCTGGCCACCGCCCAAGAGAGGAGGA 1735
Db 1917 AGTGTGCTGCTAGCTCGGGGTGCGC---GGCGGAGGAGCGAGGGTTCGAAGCGGAGGA 1973
QY 1736 TCAAGGTGCGCTTGGCGAGAGATCCCGG 1765
Db 1974 TCGCGCGCTCGCCAGGAGACGTGGCG 2003
```

RESULT 5

PCT-US02-33122-2

; Sequence 2, Application PC/TUS0233122

; GENERAL INFORMATION:

; APPLICANT: Exseed Genetics

; TITLE OF INVENTION: Starch

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: PCT/US02/33122

```
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: 60/329,525
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO. 2
; LENGTH: 2263
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (1643)
; OTHER INFORMATION: The mutation is at position 1643 in the sequence,
; OTHER INFORMATION: 1450 bp after the start codon.
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (1643)
; OTHER INFORMATION: The change is a "G" in the wild type (EX385) to an
; OTHER INFORMATION: "A" in the mutant
PCT-US02-33122-2
```

Query Match 22.4%; Score 476.4; DB 1; Length 2263;
Best Local Similarity 59.7%; Pred. No. 1.7e-56;
Matches 985; Conservative 0; Mismatches 566; Indels 99; Gaps 7;

```
QY 161 GTGCCACTTGC CGCGTGGACATCGTGTGCTGCTGCTGAGGTGCGCCCTTGTGTCACA 220
Db 408 GCGCGAGCGCGCGGATGAACGTCGTCGTCGCGCGGAGATGGCGCGTGGAGCAAGA 467
QY 221 CGGGCGCGCTGGCGATGTGACTGGTGGCTGCCTATTGAGCTGTGTCAGCGCGGCAC 280
Db 468 CCGGCGCGCTCGCGACGTCCTCGCGCGGCTGCGCGCGGCTGCGCGCGGAGCGGCGAC 527
QY 281 GGTCATGACCATTCGCTCGTACGACAGTACGCTGACGCTGGGACACCTCGGTGG 340
Db 528 GTGTATGTGCTCTCTCCCGCTAGCACAGTACAGGACGCTCGGACACCGAGCTCG 587
QY 341 TGTGTGACATCA-----TGGGCGAGAGGTCCGCTACTTCCACTCCATCAAGA 388
Db 588 TGTCCGAGATCAAGATGGGAGACGGGTACGAGACGCTAGGTTCTTCCACTGTACAAGC 647
QY 389 AGGGCGTGACCGCTGTGATGACACCGCTGCTTCTGCGCAAGTCTGGGCAAGA 448
Db 648 GCGGAGTGACCGCGTGTTCGTTGACCAACCGCTGTTCTCTGGAGAGGTTTGGGAAAGA 707
QY 449 CGGCTTCCAAAGTGTACGGCCCGCTCGGCGCTGACTACCTGGACAAACCAAGCGCT 508
Db 708 CCGAGGAGAAGATCTACGGGCGCTGCTGCTGGAACGACTACAGGAGACACCGCTCGGT 767
QY 509 TCGCCCTGTTCTGAAGCGCGCTATTGAGGTGCGCGCGTGTGCTTCCGCTTCGGC----- 561
Db 768 TCAGCTGTATGCCAGGACGACITGAAGCTCAAGGATCTCTGAGCTCAACAACAAC 827
QY 562 -----CCCGCGAGGACTGCGTCTTCGTGGCCAAAGCTGCGACTCG 604
Db 828 CATACTTCTCGGACCAATAGGGGAGGACGTCGTGTTCTGCTCAACGACTTGCACACCG 887
QY 605 CCCTGGTCCCGCTCTGCTGAAGGACGAGTACAGCCCAAGGCCAGTTCACCAAGGCA 664
Db 888 GCGCTCTCTGCTACCTCAAGAGCACTACCAAGTCCCGACGCTACAGGAGCGCA 947
QY 665 AGTGGTGTGCTATCCACACATCGCTTCCAGGCGCGATGTGGGAGGAGGCTTCA 724
Db 948 AGACCGCTTTCGCAATCCACAATCTCTACAGGCGCGGTTCGCTTCTCCGACTAC 1007
QY 725 AGGACAGAGCTGCCCGACGCGCTTTTGACAAGCTGCGCTTCTCGGACGCTATGCA 784
Db 1008 CGAGCTGAACCTCCCGGAGAGATTCAAGTGTCTTCTGATTTCTATPCGACGCTACGAGA 1067
QY 785 AGGTTTACACTGAGGCGACCCCATGGAGGAGGACGAGAAAGCCCGCTGACGGGAAGA 844
Db 1068 AG-----CCGTTGGAAG-----CCGTTGGAAG----- 1079
```

```
QY 845 CTTACAAGAAGATCAACTGGGTGAAGGGTGGCAATTATCGCCCGCAGCAAGCTGGTGAC 904
Db 1080 ---GCCGGAAGATCAACTGGATGAAGCGCGGATCTCGAGCGCGACAGGGTCTCACCG 1136
QY 905 TGTGCGCCCAAGTACCGGACCGAGATCGTGCCGATGCGCCCGCGCGGTGGAGCTGGACA 964
Db 1137 TCAGCCCTTACTACCGGAGGAGCTCATCTCCGGATCGCCAGGGGCTCCGAGCTCGACA 1196
QY 965 CCGTCATCGCCGCCAAGGCAATGTGAACGGCATGTGAACGGCATGGAGCTGGAGTGA 1024
Db 1197 ACAATCATGGCTTACCGGCAATCAGCGCATCGTCAACGGCATGAGCTCAGGAGTGG 1256
QY 1025 ACCCAAGACCGACAAGTTCCTGTCTGGCCCTACGACAGCAAGACAGCTTACGCGGCA 1084
Db 1257 ACCCAAGACCGACAAGTACATCGCGCTGAAGTACGACGTGCGACGCGGTGGAGCCA 1316
QY 1085 AGCGCGCCGCAAGAGGCGCTGCGAGCGGAGCTGGCTGCTGGAGCCCGCCAGCGCC 1144
Db 1317 AGCGCGTGAACAGAGGCGCTGCGAGGAGGCTGGGCTCGCGGTGGAGCGGAACATCC 1376
QY 1145 CCGTGTGCGCTTCAATCGCGCGCTGGAGGAGCAGAAGGGTGTGGACATCATCTCGCGC 1204
Db 1377 CGCTGGTGGCTTCAATCGCGAGCTGGAAGACAGAGGCGCCGACGTATCGGCGCG 1436
QY 1205 CCCTGCCCAAGATCCTGGCCACCCCGAAG---GTGAGATCGCCATTCCTGGGTACGGCA 1261
Db 1437 CCATCCCGCAGCTCATGAGATGTGTGAGGAGCTGCAGATCTTCTGTGGGACGGGCA 1496
QY 1262 AGCGCGCTTACGAGAGCTGTGAAGCCATCGGCACCAAGTACAGAGGCGCGCCAGG 1321
Db 1497 AGAAGAGTTCGAGCGCATGCTCATGAGCGCGGAGGAGATTCGACGACAGAGTGGCGG 1556
QY 1322 CGGTGGTCAAGTTCCTGGCGCGCTGGCGCACATGCTACCGCGCGCGCGCTTCATGC 1381
Db 1557 CGGTGGTCAAGTTCGAGCGCGCTGGCGCACACATCATGCGCGCGCGCGCTGCTCG 1616
QY 1382 TGGTGGCTCGGCTTCGAGCCCTCGGCGCTGATCAGCTGCAGCGCATGACAGGTA 1441
Db 1617 CGGTGACCAAGCGCTTCGAGCGCTGAGCGCTCATCCAGCTGAGGGGATGCGATAGGAA 1676
QY 1442 CGGTGCGCTGTAGCTTCACCGCGCGCTGGTGCACACCGTCAAGAGGCGGTCAACG 1501
Db 1677 CGCGCTGCGCTGCGCGTCCACCGGTGAGTGTGACATCATCGAAGCAAGACCG 1736
QY 1502 GCTTCCATATGGCGCTGACCGCG-----ACAAGCTGAAGAGGCTGACGCGAGC 1555
Db 1737 GGTTCACATGGCGCGCTCAGCGTCAAGCTGTAAGCTGCTGGAGCGCGGAGCTCAAGA 1796
QY 1556 CCTGGCGCCACCGTGGCGCTGCGAGCGAGGTGTTGCGGGCGCGCTACCGCGAGA 1615
Db 1797 AGGTGGCCACCATGTCAGCGCGGCTCAAGGTGCTGGCAGCGCGGCTACGAGGAGA 1856
QY 1616 TGGTGGCCAACTGCACTAGCGAGGCTGCTCTGCTGCTCAAGCGCGCGGCAAGAGTGGAGG 1675
Db 1857 TGGTGGAGACTGATGATCAGGATCTCTCTGGAAGGCGCTGCCAAGACTGGAGA 1916
QY 1676 GCGTGTGAGGAGGTGTTGATGAGCAAGGCGGCTGCGCCACCGCGCAAGAGAGGAGA 1735
Db 1917 AGTGTGTCTGATCGCTCGGGGTGCGC---GGCGGCGAGCGAGGCTCGAAGCGAGGAGA 1973
QY 1736 TCAAGTGGCGCTGTCGCGAGAGATCCCGG 1765
Db 1974 TCGCGCGCTCGCCAGGAGAGCTGGCGG 2003
```

RESULT 6

US-10-272-291-2

```
; Sequence 2, Application US/10272291
; GENERAL INFORMATION:
; APPLICANT: ExSeed Genetics
; TITLE OF INVENTION: starch
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/272,291
```

```
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: 60/329,525
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2263
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: mutation
; LOCATION: (1643)
; OTHER INFORMATION: The mutation is at position 1643 in the sequence,
; OTHER INFORMATION: 1450 bp after the start codon.
; NAME/KEY: mutation
; LOCATION: (1643)
; OTHER INFORMATION: The change is a "G" in the wild type (EX385) to an
; OTHER INFORMATION: "A" in the mutant
US-10-272-291-2

Query Match 22.4%; Score 476.4; DB 9; Length 2263;
Best Local Similarity 59.7%; Pred. No. 1.7e-56;
Matches 985; Conservative 0; Mismatches 566; Indels 99; Gaps 7;

QY 161 GTCCCACTTGGCGCTGGACATGTGATGTTGCTGCTGAGTTCGCCCTTGGTCCAAGA 220
Db 408 GCGCCAGCGCGGCATGAACGTCTTCTGTCGGCGCGAGATGGCGGTGGAGCAAGA 467
QY 221 CGGGCGGCTTGGCGGATGTGACTGGTGGCTGCTGCTATTGAGTGGTCAAGCGGGCCACC 280
Db 468 CGGGCGGCTTGGCGGATGTGACTGGTGGCTGCTGCTATTGAGTGGTCAAGCGGGCCACC 527
QY 281 GCGTCATGACATGTCGCCCTCGCTACGACCAATACCTGACGCTGGGACACCTCGGTGG 340
Db 528 GTGTCATGTCGCTCTCTCCCGCTACGACCAATACCTGACGCTGGGACACCTCGGTGG 587
QY 341 TGTGGAATCA-----TGGGCGAGAAGTCCGCTACTTCCACTCCATCAAGA 388
Db 588 TGTCCGAGATCAAGATGGGAGCGGTAGGACGCGTACGACGCTGAGTTCCTCCACTGTACAAGC 647
QY 389 AGGCGCTGACACCGCTGTGATTTGACACCGCTGTTCTTGGGCAAGGTCTGGGCAAGA 448
Db 648 GCGAGTGGACCGCTGTTGCTGTGACCACTGTTCTGAGAGAGGTTTGGGGAAGA 707
QY 449 CGGCTTCCAAAGCTGTACGGCGCGCTCGGCGCTGACTACTGSAACCAACAGCGCT 508
Db 708 CCGAGGAGAAGATCTACGGCGCTGTCGCTGGAACGGACTACAGGAGCAACCAAGCTCGGT 767
QY 509 TCGCCCTGTCTGCAAGCGCGCTATTGAGCTGCCCGGTGCTGCGCTTCCGCTTCCGCTTCC 561
Db 768 TCAGCTGTCTATGCCAGGAGCTTTGAAGCTTCAAGGATCCTTGAAGCTTCAACCAAGC 827
QY 562 -----CCGCGGAGGAGTGGCTCTTCTGTCGCAAGCTGGCTTCTCGGAGCGGTATCC 604
Db 828 CATACTTCTCGGACCAATAGGGGAGGAGCTGCTGTTGCTGCTGCAAGCTGCAACAGC 887
QY 605 CCGTGTGCGCGCTCTGCTGAAGAGAGTACACGCCCAAGGCGGCTTCAACAGGCGCA 664
Db 888 GCGCTCTCTGCTACCTCAAGAGCAACTACCACTCCAGGCACTTACAGGAGCAACCAAG 947
QY 665 AGTGTGTCTGCTATCCCAACATCGCTTTCAGGGCGGCTGCTGGAGGAGGCTTCA 724
Db 948 AGACCGCTTCTGCTATCCCAACATCTCTACGAGGCGGCTTTCGCTTCTCGGAGTACC 1007
QY 725 AGGACACGAAGCTGCCCGCGCTTTTGAAGCTGGCTTCTCGGAGCGGTATGCCA 784
Db 1008 CCGAGCTGAACCTCCCGGAGAGATTCAAGTCTGCTTTCGATTTTCATCGAGCTACGAGA 1067
QY 785 AGTTTACACTGAGGCCACCCCGATGAGGAGGAGACGAGAGAGCCCGCTGACGGAGAA 844
Db 1068 AG-----CCGCGGAG-----CCGCGGAG----- 1079
```

```
QY 845 CCTACAGAGATCAACTGGCTGAAGGGTGGCATTTATCGCCCGCCGACAGCTGGTACTG 904
Db 1080 ---GCCGGAGATCAACTGGATGAAGCCGGGATCTCTCGAGCGCGACAGGCTCTCACCG 1136
QY 905 TGTTCGCCCAACTACGCGACGAGATCGCTGCCGATCGCGCGCGGTGTGGAGCTGGACA 964
Db 1137 TCAGCCCCCTACTACGCCGAGGAGCTCATCTCCGGCATCGCAGGCGCTGGAGCTCGACA 1196
QY 965 CGGTCACTCCGCGCCAAAGGCAATFAGGGCATTTGAAACGGCATFAGACATFAGAGAGTGA 1024
Db 1197 ACATCATGGCGCTCACCGCATCACCGCATCGTCAACGSCATGGACGTGAGGAGTGGG 1256
QY 1025 ACCCAAGACGACAGATTCCTGTCTGCGCCCTTACACACAGACACGCTCTAGCCCGGA 1084
Db 1257 ACCCAAGACGACAGATTCCTGTCTGCGCCCTTACACACAGACACGCTCTAGCCCGGA 1316
QY 1085 AGGCGCCGCCAAAGAGGCGCTGACGCGGAGTGGCGCTGCTGTGGACCCCAACCGCC 1144
Db 1317 AGGCGCTGAACAAGGAGGCGCTGACGCGGAGTGGCGCTGCTGTGGACCCCAACCGCC 1376
QY 1145 CCTGTTCCTTCATCGCGCGCTGAGAGAGAGAGAGGCTGGGACATCATCTGGCCG 1204
Db 1377 CGTGTGTGCTTCATTCGCGAGCTGGAAGAGAGAGAGGCGCCCGACGTCATGGCGCGC 1436
QY 1205 CCTGTGCCAAGATCTCTGGCCACCCCAAG---GTGCAGATCGCCATCTTGGGTACCGGA 1261
Db 1437 CCATCCCGAGCTCATGGAGATGTTGGAGAGCTGCAGATCGTTCGTCTGGGACCGGGA 1496
QY 1262 AGGCGCCCTACGAGAGCTGGTGAAGCCCATCGCCACCAAGGAGAGGCGCGCAAG 1321
Db 1497 AGAAGAAGTTCGAGGCGCATGCTCATGAGCGCGGAGGAGAGTTCACGCAAGTTCGCG 1556
QY 1322 GGTGTCAAGTTCGCGCGCGCTGCGGCACATGCTACCGCGCGCGCGCTCATGTC 1381
Db 1557 CGTGTGCTCAAGTTCACAGCGCGCTGCGGCACCATCATGCGCGCGCGCGCTGCTG 1616
QY 1382 TGTGTGCTTCGCGCTTCGAGCGCTGCGCGCTGATCCAGCTGCACGCCATGCACTAGGTA 1441
Db 1617 CGTCAACAGCGCTTCGAGCGCTTCGAGCGCTCATCCAGCTGCAGGGGATGCGATCGGA 1676
QY 1442 CGTGTGCTTCGCTTACCGCGCGCTGCTGACACCGCTCAAGGAGGCGTCAACCG 1501
Db 1677 CGCGCTGCGCTGCGCGCTCACCGGTGACTGCTGCACACCATCATCGAAGCGAAGCG 1736
QY 1502 GCTTCCATGAGGCGCGCTTACCGCG-----ACAAGTGGACAGGCTGACCGCAG 1555
Db 1737 GGTTCACATGGCGCGCTTCAGCGTCACTGTAAAGTTCGTGAGCGCGCGCGTCAAGA 1796
QY 1556 CCTGTGCGCCACCGTTCGCGCGCTGCGAGGAGTGTTCGCGCGCGCTTACCGCGGA 1615
Db 1797 AGTGTGCCACCATTCGAGCGCGCTCAAGGTGTGCGAGCGCGCGCTTACGAGGAGA 1856
QY 1616 TGTGTGCCAACTGCATFAGCCAGGACCTTCTCTGTGTCCTCAAGCGCGCGCGCGAGG 1675
Db 1857 TGTGTGAGGAATGCATGATCCAGGATCTCTCTGTGGAAGGCGCTTCCCAAGACTGGAGA 1916
QY 1676 GCTGCTGAGAGGAGTGTGTACGGCAGGCGCGCTGCGCGCGCGCGCGAGAGGAGA 1735
Db 1917 AGTGTGCTGTAGCGCTCGCGCGTGC-----GGCGCGCGCGCGAGGCTGAGAGGAGA 1973
QY 1736 TCAAGTGCCTGTGCGGAGAGATCCCGC 1765
Db 1974 TCGCGCGCTCGCAAGGAGAGAGTGGCCG 2003
```

RESULT 7

US-10-336-753-46

; Sequence 46, Application US/10336753
; GENERAL INFORMATION:
; APPLICANT: Guan, Hanping
; APPLICANT: Keeling, Peter L.
; TITLE OF INVENTION: PLANT LIKE STARCHES AND THE METHOD OF MAKING THEM IN
; TITLE OF INVENTION: HOSTS

```
; FILE REFERENCE: 2461-52  
; CURRENT APPLICATION NUMBER: US/10/336,753  
; CURRENT FILING DATE: 2003-01-06  
; PRIOR APPLICATION NUMBER: US/09/402,254  
; PRIOR FILING DATE: 1999-10-01  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US98/06660  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-03  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/042,939  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-04-04  
; NUMBER OF SEQ IDS: 77  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 46  
; LENGTH: 1915  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (join(1..1815, 1819..1914))  
; US-10-336-753-46
```

```
Query Match 22.3%; Score 472.8; DB 9; Length 1915;  
Best Local Similarity 59.6%; Pred. No. 5.2e-56;  
Matches 982; Conservative 0; Mismatches 567; Indels 99; Gaps 7;  
  
QY 163 GCCACTTGGCGCTGGACATCGTGTGCTGTGAGTGGCCCTTGGTCCAGACG 222  
Db 217 GCCAGCGCGCGCATGAACGCTGCTTCGTCGGCGCGAGATGGCGCGTGGAGCAAGCC 276  
QY 223 GGGCGCTGGCGCTGATGACTGGCTGCTCTTATGAGTGAAGCGCGCCACCGC 282  
Db 277 GGGCGCTGGCGCTGCTGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 336  
QY 283 GTCATGACCATTCGCCCTTCGCTACGACCATGAGCTGACCGCTGGGACACCTCGGTGTC 342  
Db 337 GTCATGCTGCTCTCTCCCGCTACGACCATGACAGGAGCGCTGGGACACCGCGTGTG 396  
QY 343 GTGGACATCA-----TGGGGGAGAAGGTCGGCTTCTTCCACTGCTACAGCGC 456  
Db 397 TCCGAGATCAAGATGGGAGACGGGTAGGAGCGGTCTTCCACTGCTACAGCGC 450  
QY 391 GGGCTCACCGCGTGTGGATTGACACCGCTTCTGCGCGAGGTCTGGGCAAGAC 450  
Db 457 GAGTGGACCGCGTGTGCTGACACCGCTTCTGCGAGAGGTTGGGAAAGAC 516  
QY 451 GGTCTCAAGTGTACGCGCGCGCTGCGCGCTGAGTACCTGACACCAACAGCGCTTC 510  
Db 517 GAGGAGAAGATCTACGCGCGCTGCGCTGGAACGAGTACAGGAGCAACAGTGGGTTTC 576  
QY 511 GCCTGTCTGCAAGCGCGCTATTGAGGCTGCGCGGTGCTGCGCTTGGCC----- 562  
Db 577 AGCTGCTATGCCAGGAGCACTTGAAGCTTCAAGGATCCTGAGCTTCAACACACCA 636  
QY 563 -----CCGCGAGGAGTGGCTTCTGTCGCGCAACAGTGGCACTCGCC 606  
Db 637 TACTTCTCGGACCATACGCGGAGGAGTGTGCTGTCGCAACAGCTGGCAACCGCG 696  
QY 607 CTGGTCCCGTCTCTGTGAAGGACGAGTACAGCGCGCGAGGCGAGTTCACAGGCGCAAG 666  
Db 697 CTTCTCTGCTGTACCTCAAGACCACTACACAGTCCACAGGCTTACAGGAGCGCAAG 756  
QY 667 TGGTGTGCTGTATCCACAACATCGCTTCCAGGCGCGCATGTGGGAGGAGGCTTCAAG 726  
Db 757 ACCGCTTCTGCTATCCACAACATCTCTACAGGCGCGGTTCGCTTCTCCGACTACCG 816  
QY 727 GACACGAAGTCCCGCGAGCGCGCTTTGACAGCTGCGCTTCTCGACGCGCTATGCCAAG 786  
Db 817 GAGCTGAACCTCCCGGAGAGATTCAGTCTGCTCTTCGATTTCAATCGCGCTACGAGAAG 876  
QY 787 GTTTACACTGAGGCCACCGCGGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 846  
Db 877 -----CCGCTGGAG----- 886  
QY 847 TACAAGAGATCAACTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
```


Db 2377 TATATGTGTTCTACAGATTTGACCTGTGTGTTCTCATTCAGTTTACAAAGCTATCGCTA 2436
QY 1437 CGGTACCGTGCCTGTAGCTCCACCGGGGCTGTGTCGACACCGTCAAGAGGGGT 1496
Db 2437 TGGATCTGACCAATTTGCTCAACAGGTGATATGTTGACATGTCGAAGAGGCTT 2496
QY 1497 CACCGGTTCCACATGGGCGCCTGA-----ACCCGACAAAGCTGGAGGCTGACGC 1550
Db 2497 CACTGGATTCAGATGGTGCCTTCAATGTTGAATGTGATGCTGGATCCGGCTGATG 2556
QY 1551 CGACGCCCTGGCGCCACCGTGCCTGCGGTCGCGACAGAGTGTGTCGGGGCGGCTACCC 1610
Db 2557 GGATGCTATATCAAGAGCTGTCAAAAGGGCCCTTCAGTCTATGGAACCTCCAGCTTTAC 2616
QY 1611 CGAGATGTGGGCAACTGATCAGCAGGACCTGCTCGTGTCCAAAGCCGCGCAGAAGTG 1670
Db 2617 AGAATATCAAGAACTGCATGGCTCAAGATCTTTTCATGGAAGGGGCTGCTAAGAGTG 2676
QY 1671 GGAGGCGCTGCTG 1683
Db 2677 GGAGGAAGTGTG 2689

RESULT 15

US-10-336-753-48
; Sequence 48, Application US/10336753
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter L.
; TITLE OF INVENTION: PLANT LIKE STARCHES AND THE METHOD OF MAKING THEM IN
; TITLE OF INVENTION: HOSTS
; FILE REFERENCE: 2461-52
; CURRENT APPLICATION NUMBER: US/10/336,753
; PRIOR FILING DATE: 2003-01-06
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US98/06660
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/042,939
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-04-04
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 2423
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (Join(1..2094, 2098..2103, 2107..2304, 2308..2421))
US-10-336-753-48

Query Match 10.3%; Score 219.2; DB 9; Length 2423;
Best Local Similarity 50.6%; Pred. No. 2.3e-21;
Matches 801; Conservative 0; Mismatches 673; Indels 108; Gaps 7;
QY 144 TGTACTGTCACATGTCGCACCTGTCGCGCTGACATCGATGTTGCTGCTGAGGT 203
Db 588 TGAACCTGGCCCTTTGGCTGGCCCTTAATGTGATGAACCTGCTGCTGGTCTCTGAATG 647
QY 204 CGCCCTTGGTCCAAAGAGGGGGGCTGGGCGATGTGACTGCTGGCTCGCTATTGAGCT 263
Db 648 TGTCTCTTCTCAAGACAGGTGGCTTGGAGATGTCGTTGGGTGCTTTGCCCTAAGGCTCT 707
QY 264 GGTCAAGCGCGCCACCGGCTCATGACCATTCGCCCTCGCTACGACCACTAGCTGACGC 323
Db 708 GCGGAGGAGGACACCGCTGTTATGTCGTGATACCAAGATATGAGAGTATGCCGAAGC 767
QY 324 CTGGGACACCTCGTGGCTGTCGACATCATGCGCGCAGAGGTCCGCTACTTCCACTCAT 383
Db 768 CCGGG--ATTTAGTGTATAGAGACGTTTACAAAGGTAGCTGGAGAGTTCAGAAAGTTACT 825
QY 384 CAAGAAGGCGTGCACCGGCTGTGGATTGACCAACCCCTGTTCTCTGGCCAAAGTCTGGG 443

Db 826 TATTTTACTCTTACATTTGATGGAGTTGATTTTGTATTCGTAGAACCCCTCCTCTCCCG 885
QY 444 CAAGACCGGCTCAAGCTGTACGGCCCGCTCCGGCTGACTACTGTGACAAACACAA 503
Db 886 CAC-----CGCACAATAATATTTATGGGGAGAAGATTTGGATTTTGA 932
QY 504 GGGCTTGCCTTGTTCGAAGSCCGCTATTAGGCTGCCCGCTGCTGCCCTTCCGCGCC 563
Db 933 GCGCATGATTTTGTCTGCAAGSCCGCTGTGAGGTTCCATGATGCTCCATGTGGCG 992
QY 564 -----CGGCGAGGACTGCTCTTCGTTGGCCACAGCTGACATCCGCCCTGT 611
Db 993 TACTGTCTATGTGATGGCAACTTAGTTTTCATGTTGTAATGATTTGACATACCGCACTT 1052
QY 612 GCGCGTCTGCTGAAGAGCAGGTACCAAGCCCAAGGCGCAGTTTCAACCAAGGCCAAGTCGGT 671
Db 1053 GCTGTCTATCTAAAGSCCTATTACCGGACATATGTTGATGCAATGCTGCTGCTCTGT 1112
QY 672 GTTGGCTATCCACATCGCTTCCAGGCGCGCATGTGGAGAGGCTTTCGAAGACAC 731
Db 1113 GCTGTGATACACATTTGCTCATCAGGCTGCTGGCCCTGTAGACGACTTCGTCAATTT 1172
QY 732 GAAGCTGCCCGGCGCTTTGACAAAGCTGGCTTCTCGAGGCGTATGCCAAGCTTTA 791
Db 1173 TGACTTGGCTGAACACTATACGACCACTTCAAACTGTATGACAACTTTGGTGGGATCA 1232
QY 792 CACTGAGGCCACCCCATGGAGGAGGACGAGAAGCCCGCTGACGGGAAGACCTTACAA 851
Db 1233 C----- 1233
QY 852 GAAGATCACTGGCTGAAGGTGGCATTTATCGCCCGCGCACAGCTGGTACTGTGCGCC 911
Db 1234 ----AGCAACAGTTTTTGTCTCGGGGCTGAAGACGCGCAGACCGGCTGGTACCGTTAGCAA 1289
QY 912 CAACTACGCGACCGAGATCGTCCGATGCGCGCGGCTGTGAGCTGGACACCGCTCAT 971
Db 1290 TGTCTACATGTGGAGCTGAAGACTTCGGAAGCGGCTGGGCGCTCCACACATCAATA 1349
QY 972 CC---CGCCAAAGGGGCTTGAAGGCAATGTGAACGCATGGACATTTAGAGAGTGGACCC 1028
Db 1350 CCAGAACGACTGGAAGCTGCAGGCGCATCGTGAACGCGCATCGACATGAGCGAGTGAACCC 1409
QY 1029 CAAGACCGACAAGTTCTGCTGCGCCCTACGACACAGACAGC-----GTCATA 1076
Db 1410 CGCTGTGAGCTGSCACTCCACTCCGACGACTACACCACTACACTTCGAGACGCTGA 1469
QY 1077 CGCGGCAAGCGCGCGCCCAAGGAGCGCTGCAAGCGCGAGCTGGGCTGCTGTGACCC 1136
Db 1470 CACCGCAAGCGGCACTGCAAGCGCCCTGCAAGCGCGAGCTGGGCTGCGAGTCCGCGA 1529
QY 1137 CACCGCCCGCTGTTGCGCTTTCATCGCGCGCTTGAAGAGCAGAGAGGTTGAGACATCAT 1196
Db 1530 CGACGTGCCACTGATCGGGTTCATCGGGGCGCTGGACACCAAGAGGGGTGACATCAT 1589
QY 1197 CTGCGCGCGCTGCCAAAGATCCTGCGCCACCCCAAGGTGACATGCCATCCTGGGTAC 1256
Db 1590 CGCGGACGCGATCCACTGGATC---GCGGGGAGGAGCTGACGCTGCTGATGCTGGCAC 1646
QY 1257 CGCAAGCGCGCTACGAGAAGCTGCTGAACGCGCATCGGCGCACCAAGTACAAAGGCGCGCG 1316
Db 1647 CGGCGGCGGCGACTGAGGACATGCTCGCGCGTTCGAGTCGAGCACACAGCGACAGGT 1706
QY 1317 CAAGGCGTGTCAAGTTCTCGGCGCGCTTGGCGCACATGCTACCGCGCGGCGCGACAT 1376
Db 1707 GCGCGGTGGTGGGTTCTCGGTGCGCCCTGGCGCACCGCATCACGCGGCGCGGACAT 1766
QY 1377 CATGCTGTGCGCTCGCGCTTCGAGCGCTGCGGCTGATCCAGCTGACGCGCATGCACTA 1436
Db 1767 CTGCTGTGTCGCGTTCGAGCGCTGCGGGCTGAACAGCTCTACGCCATGGCGTA 1826
QY 1437 CGGTACGCTGCGCGTGTAGCTTCCACCGCGCGCTTGGTCGACCGCTCAAGAGGCGGT 1496
Db 1827 CGGACCGTGGCGTGTGACCGCGTGGGCGGCTCCGCGACACGCTGGCGCGTTCGA 1886

QY	1497	CACCGGCTTCACATGGCGCCCTGACCCCGACAGCTGGACGAGGCTGACGCGGACGC	1556
Db	1887	CCCGTTCAAGACACACCGGCTCGGGTGGACGTTTCGACCGCGGAGCGGAACGGATGAT	1946
QY	1557	CCTGGCGGCACCGTGGCGGTGCCAGCGAGGTGTTTGGGGCGCGCTACCCGAGAT	1616
Db	1947	CGACGGCTCTCGCACTGCTCACACAGTACCGGACTCAAGGAGAGCTGGCGCGCTG	2006
QY	1617	GGTGGCCAACTGCATCAGCCAGGACCTGTCTGTCTCCAGCCCGCCAGAAAGTGGAGGG	1676
Db	2007	CAGGGCGCGCGGATGGCGAGGACCTCAGCTGGGACCGCCGCGCTGTGTATGAGGA	2066
QY	1677	CCTGCTGGAGGAGTGTGTAC	1698
Db	2067	CGTGTCTGTCAAGGCGAAGTAC	2088

Search completed: June 7, 2003, 02:49:00
 Job time : 1012.93 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:49:45 ; Search time 37.5509 Seconds
(without alignments)
2512.365 Million cell updates/sec

Title: US-09-980-771A-3
Perfect score: 3641
Sequence: 1 MAVASTSPSSARPVINA.....SASKTSAKPLVSAATPKSA 708

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1450	39.8	606	13	AA25476		Granule-bound star
2	1426	39.2	609	19	AAW56485		Oryza sativa starch
3	1405.5	38.6	563	21	AAW04667		Arabidopsis thalia
4	1405.5	38.6	610	21	AAW04666		Arabidopsis thalia
5	1405.5	38.6	610	23	ABW91229		Herbicidally activ
6	1390	38.2	527	21	AAW04668		Herbicidally activ
7	1360.5	37.4	637	19	AAW16604		Protein encoded by
8	1358.5	37.3	533	19	AAW56484		Zea mays waxy gene
9	1341.5	36.8	502	13	AA25474		Central fragment f
10	1194	32.8	238	22	AAB31175		A granule bound st

11	904	24.8	534	21	AAW49307		Wheat starch synth
12	900.5	24.7	647	20	AAW09004		Wheat starch synth
13	898	24.7	792	23	ABW92160		Herbicidally activ
14	890	24.4	756	21	AAW50818		Wheat soluble star
15	885.5	24.3	649	19	AAW58218		Maize starch synth
16	884.5	24.3	671	19	AAW23937		Wheat soluble star
17	873.5	24.0	626	15	AAW51231		Soluble rice starch
18	872	23.9	652	23	ABW93595		Herbicidally activ
19	869.5	23.9	583	19	AAW56488		Zea mays soluble s
20	865.5	23.8	539	19	AAW56491		Zea mays starch sy
21	860.5	23.6	583	19	AAW70894		Maize starch solub
22	855.5	23.5	804	19	AAW70892		Maize starch solub
23	848	23.3	802	23	AAW99845		Modified barley st
24	848	23.3	813	23	AAW99844		Barley cultivar Mo
25	847.5	23.3	698	19	AAW56487		Zea mays soluble s
26	847	23.3	812	23	AAW99847		Barley line MK6827
27	845	23.2	798	21	AAW37566		Wheat starch synth
28	843	23.2	799	21	AAW37567		Wheat starch synth
29	842	23.2	799	21	AAW37597		Wheat starch synth
30	842	23.1	799	19	AAW23938		Wheat starch synth
31	826.5	22.7	812	23	AAW99846		Wheat granule-boun
32	826	22.7	597	21	AAW37568		Wheat starch synth
33	767	21.1	456	23	AAW51865		Rice starch synth
34	726	19.9	459	17	AAW99540		Soluble starch syn
35	702.5	19.3	669	19	AAW70893		Maize starch solub
36	702.5	19.3	669	19	AAW56486		Zea mays soluble s
37	669	18.4	483	23	ABW54012		Lactococcus lactis
38	647.5	17.8	476	23	ABW27592		Streptococcus poly
39	580	15.9	477	13	AAW25462		giga. Escherichia
40	580	15.9	477	13	AAW53890		E. coli glycochen s
41	570.5	15.7	677	17	AAW99539		Soluble starch syn
42	570.5	15.7	1230	18	AAW17785		Potato tuber solub
43	570.5	15.7	1230	21	AAW49306		Potato starch synth
44	564.5	15.5	1025	23	ABW90967		Herbicidally activ
45	559	15.4	735	18	AAW22729		Z. mays starch syn

ALIGNMENTS

RESULT 1
AA25476
ID AAR25476 standard; Protein; 606 AA.
XX AAR25476;
AC AC
DT DT
DT 15-JAN-1993 (first entry)
XX
DE Granule-bound starch synthase of potato.
KW GBSS gene; amylopectin; amylose production;
KW inhibition; promoter; antisense construct.
XX
OS Solanum tuberosum.
XX
XX WO9211376-A.
PN
PD 09-JUL-1992.
XX
PF 20-DEC-1991; 91WO-SE00892.
XX
PR 21-DEC-1990; 90SE-0004096.
XX
PA (AMYL-) AMYLOGENE HB.
XX
PI Hofvander P, Persson PT, Tallberg A, Wikstrom O;
XX
DR WPI; 1992-250096/30.
DR N-PSDB; AAQ26404.
XX
PT Modifying potatoes to form amylopectin starch - using an
PT anti-sense construct to inhibit granule-bound starch synthase

Db 475 ADVLAVPSRFEPCCGLIQLQGMRYGTPPCACATGGGLVDTVIRGKTGFHMGRLSSVDCKVVEP 534
 Qy 515 ADADALAATVRRASEVAGGRYPPEMVANCISQDLSWSKPAOKWEGLLVEVYGRGG-----570
 Db 535 SDVKKVAATLAKRAIKVYGTTPAYEWRVRCNMQDLSWSKGPARNWNLV--LGLGVAGSAPG 592
 Qy 571 -----VATAKKEETKVP 582
 Db 593 IEGDEIAPLAKENVAAP 609
 RESULT 3
 AAG04667
 ID AAG04667 standard; Protein; 563 AA.
 XX AC
 XX AC AAG04667;
 XX DT 17-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 781.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX FN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126284.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.

09-AUG-1999; 99US-0147935.
 10-AUG-1999; 99US-0148171.
 11-AUG-1999; 99US-0148319.
 12-AUG-1999; 99US-0148341.
 13-AUG-1999; 99US-0148565.
 13-AUG-1999; 99US-0148684.
 16-AUG-1999; 99US-0149368.
 17-AUG-1999; 99US-0149175.
 18-AUG-1999; 99US-0149426.
 20-AUG-1999; 99US-0149722.
 20-AUG-1999; 99US-0149723.
 20-AUG-1999; 99US-0149729.
 23-AUG-1999; 99US-0149902.
 23-AUG-1999; 99US-0149930.
 25-AUG-1999; 99US-0150566.
 26-AUG-1999; 99US-0151065.
 27-AUG-1999; 99US-0151066.
 27-AUG-1999; 99US-0151080.
 30-AUG-1999; 99US-0151303.
 31-AUG-1999; 99US-0151438.
 01-SEP-1999; 99US-0151930.
 07-SEP-1999; 99US-0152363.
 10-SEP-1999; 99US-0153070.
 13-SEP-1999; 99US-0153758.
 15-SEP-1999; 99US-0154018.
 16-SEP-1999; 99US-0154039.
 20-SEP-1999; 99US-0154779.
 22-SEP-1999; 99US-0155139.
 23-SEP-1999; 99US-0155486.
 24-SEP-1999; 99US-0155659.
 28-SEP-1999; 99US-0156458.
 29-SEP-1999; 99US-0156596.
 04-OCT-1999; 99US-0157117.
 05-OCT-1999; 99US-0157753.
 06-OCT-1999; 99US-0157865.
 07-OCT-1999; 99US-0158029.
 08-OCT-1999; 99US-0158232.
 12-OCT-1999; 99US-0158369.
 12-OCT-1999; 99US-0159293.
 13-OCT-1999; 99US-0159293.
 13-OCT-1999; 99US-0159294.
 13-OCT-1999; 99US-0159295.
 14-OCT-1999; 99US-0159329.
 14-OCT-1999; 99US-0159330.
 14-OCT-1999; 99US-0159331.
 14-OCT-1999; 99US-0159637.
 14-OCT-1999; 99US-0159638.
 18-OCT-1999; 99US-0159584.
 21-OCT-1999; 99US-0160741.
 21-OCT-1999; 99US-0160767.
 21-OCT-1999; 99US-0160768.
 21-OCT-1999; 99US-0160770.
 21-OCT-1999; 99US-0160814.
 21-OCT-1999; 99US-0160815.
 21-OCT-1999; 99US-0160980.
 22-OCT-1999; 99US-0160981.
 22-OCT-1999; 99US-0160989.
 22-OCT-1999; 99US-0161404.
 25-OCT-1999; 99US-0161405.
 25-OCT-1999; 99US-0161406.
 26-OCT-1999; 99US-0161359.
 26-OCT-1999; 99US-0161360.
 26-OCT-1999; 99US-0161361.
 26-OCT-1999; 99US-0161920.
 28-OCT-1999; 99US-0161992.
 28-OCT-1999; 99US-0161993.
 29-OCT-1999; 99US-0162142.

Query Match 38.68; Score 1405.5; DB 21; Length 563;
 Best Local Similarity 49.68; Pred. No. 3.7e-102;
 Matches 299; Conservative 76; Mismatches 163; Indels 65; Gaps 14;
 3 VASTSRSSRARPPIVINAASFVCKTANQLLRLRLARSARKTSRSVATGATGC--ALD 60

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140923.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

```

PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 21-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match 38.6%; Score 1405.5; DB 21; Length 610;
Best Local Similarity 49.6%; Pred. No. 4.2e-102;
Matches 299; Conservative 76; Mismatches 163; Indels 65; Gaps 14;

QY 3 VASTSRPSARPVINAASFVGGVKKTAQNLRLARGSARKSTSRSAVTGATGATC--ALD 60
DB 49 VDNLQRRSQAKPV-----SAKSSKRSKVKTAGKIVCEKGM 85

QY 61 IVMAAEVAPWSKTGGLGVTGGLPIELVKRGHRVMTIAPRYDQYADADMTSVVVDI-MG 119
DB 86 VIFIGAEVGPWSKTGGLGVTGGLPPALAAARGHRVMTICPRDYQKDAMDTCVVQIKVG 145

QY 120 EK---VRYFHSIKKGVRHWIDHPWFLAKVWGTGSKLYGPRSGADYLDNHRKRALFCKA 176
DB 146 DKVENRFFHCYKRGVDRVFDHPFLAKVWGTGSKLYGPRSGADYLDNHRKRALFCKA 205

QY 177 ATEARVLPF-----GP-GEDCVFVANDHSAVLPVLLKDEYQPKGQFTKAKSVLAIH 228
DB 206 ALEAPQVNLNLSKSYFSGPYGDEVFVANDHSAVLPVLLKDEYQPKGQFTKAKSVLAIH 265

QY 229 NIAFOGRMWEAFKDTKLP---QAFDKLAFSDGYAKVYTEATPMEDEKPLITGKTYKK 285
DB 266 NIAYQGRFAFDYSLINLPISFKSFD---FMDGYEK-----PVKG---RK 305

QY 286 INWLKGGIITAADKLIVTSNRYATEATAAAGGVGLDTVIRAKGIEGIVNGMDEEWNPKT 345
DB 306 INWMAAILEARHVLTVSPYIAQELISGVDGRGVELHLYLKMKTIVSGIINGMDQWNPST 365

QY 346 DRFLSAPYQDQSVYAGKAALQALQELGLPVDPTAPLFAFGRLEEQKGVDTIILAALPK 405
DB 366 DKYIDIKYDITVDKAPLKEALQAAVGLPVDROPVTVIGFGRLEEQKGVDTIILAALPK 425

QY 406 ILATPKVQJAILGTGKAAYEKLVNAIGTYKGRAGKVVKFSAPLAHMLTAGADFMVLPSP 465
DB 426 FMGL-NVOMVILGTGKKMEAGILEEKEFPKGKAVGAKFNVPVLAHMITAGADFTIIVPSR 484

QY 466 FEPCLGLIQLHMHYGNPVVASTGGLVDTVKGVTGFHMGALNP---DKLDEADADALAAT 523
DB 485 FEPCLGLIQLHMHYGNPVVASTGGLVDTVKGVTGFHMGALNP---DKLDEADADALAAT 544

QY 524 VRRASEVFAAGRPYEMVANCISQDLWSKPAQKWEGLLEEVYGVGKGVATAKKEETKVPV 583
DB 545 VTRAVAVYGTSAQEMVNCMQDQFSKGPRLWEKVLVLSLNAV-GSEAGTEGEETI-APL 602

QY 584 AEK 586

```

```

DB 603 AKE 605

RESULT 5
ABB91229
ID ABB91229 standard; Protein; 610 AA.
XX
AC ABB91229;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 440.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB ) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 440; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
SQ Sequence 610 AA;

Query Match 38.6%; Score 1405.5; DB 23; Length 610;
Best Local Similarity 49.6%; Pred. No. 4.2e-102;
Matches 299; Conservative 76; Mismatches 163; Indels 65; Gaps 14;

QY 3 VASTSRPSARPVINAASFVGGVKKTAQNLRLARGSARKSTSRSAVTGATGATC--ALD 60
DB 49 VDNLQRRSQAKPV-----SAKSSKRSKVKTAGKIVCEKGM 85

QY 61 IVMAAEVAPWSKTGGLGVTGGLPIELVKRGHRVMTIAPRYDQYADADMTSVVVDI-MG 119
DB 86 VIFIGAEVGPWSKTGGLGVTGGLPPALAAARGHRVMTICPRDYQKDAMDTCVVQIKVG 145

QY 120 EK---VRYFHSIKKGVRHWIDHPWFLAKVWGTGSKLYGPRSGADYLDNHRKRALFCKA 176
DB 146 DKVENRFFHCYKRGVDRVFDHPFLAKVWGTGSKLYGPRSGADYLDNHRKRALFCKA 205

QY 177 ATEARVLPF-----GP-GEDCVFVANDHSAVLPVLLKDEYQPKGQFTKAKSVLAIH 228
DB 206 ALEAPQVNLNLSKSYFSGPYGDEVFVANDHSAVLPVLLKDEYQPKGQFTKAKSVLAIH 265

QY 229 NIAFOGRMWEAFKDTKLP---QAFDKLAFSDGYAKVYTEATPMEDEKPLITGKTYKK 285
DB 266 NIAYQGRFAFDYSLINLPISFKSFD---FMDGYEK-----PVKG---RK 305

QY 286 INWLKGGIITAADKLIVTSNRYATEATAAAGGVGLDTVIRAKGIEGIVNGMDEEWNPKT 345
DB 306 INWMAAILEARHVLTVSPYIAQELISGVDGRGVELHLYLKMKTIVSGIINGMDQWNPST 365

QY 346 DRFLSAPYQDQSVYAGKAALQALQELGLPVDPTAPLFAFGRLEEQKGVDTIILAALPK 405
DB 366 DKYIDIKYDITVDKAPLKEALQAAVGLPVDROPVTVIGFGRLEEQKGVDTIILAALPK 425

QY 406 ILATPKVQJAILGTGKAAYEKLVNAIGTYKGRAGKVVKFSAPLAHMLTAGADFMVLPSP 465
DB 426 FMGL-NVOMVILGTGKKMEAGILEEKEFPKGKAVGAKFNVPVLAHMITAGADFTIIVPSR 484

QY 466 FEPCLGLIQLHMHYGNPVVASTGGLVDTVKGVTGFHMGALNP---DKLDEADADALAAT 523
DB 485 FEPCLGLIQLHMHYGNPVVASTGGLVDTVKGVTGFHMGALNP---DKLDEADADALAAT 544

QY 524 VRRASEVFAAGRPYEMVANCISQDLWSKPAQKWEGLLEEVYGVGKGVATAKKEETKVPV 583
DB 545 VTRAVAVYGTSAQEMVNCMQDQFSKGPRLWEKVLVLSLNAV-GSEAGTEGEETI-APL 602

QY 584 AEK 586

```

Db	266	NIATQGRFAFDYSLNLNLPISFKGSFD---	FMQGYEK	-----PVKGG--	-RX	305
QY	286	INWLKGGIIAADKLVTWSPNATETAADAAGGVELDTVIRAKGTEGIWNGMDIEWNP	KI	345		
Db	306	INWKAALIEAHRVLTVSPYTAQELISGDRGVELHKYLRKMTVSGIINGMDVQEWNP	ST	365		
QY	346	DKFLSAPYDQNSVYAGKAAKEALQAEGLPVPDTAPLFAFIGNLEEKGKGVDDIILAALPK	405			
Db	366	DKYIDIKYDITVTDAKPLIKEALQAAVGLPVDKRDVPIVIGFIGNLEEKGKGS	DLVEALSK	425		
QY	406	ILATPKYQIATLGTGKAAEYKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADPMLVPSR	465			
Db	426	FMGL-NVOMVLIGTKKMEQAIIIELEKFKPKAVGAKFNVP	LAHMITAGADEFLIIVPSR	484		
QY	466	FEPCGLIQLHAMHYGTVPVASTGGLVDTVKGVTFGHMGALNP--DKLDEADADALAAT	523			
Db	486	FEPCGLIQLHAMRYGTVPVASTGGLVDTVKDGVTGFHIGRFNVKCEVVD	DDVDVIATAKA	544		
QY	524	VRRASEVFAAGRYPEMVANCISQDLSKSPAKQWGLEEVVYKGGVATAKKEEIKVPV	583			
Db	545	VTRAVAVYGTSAMQEMVKNCKMDQFSNKGPARLWEKVLISLVYA-GSEAGTEGEI	-APL	602		
QY	584	AEK	586			
Db	603	AKE	605			
RESULT 6						
AAG04668						
ID	AAG04668 standard; Protein; 527 AA.					
XX	AAG04668;					
XX	17-OCT-2000 (first entry)					
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 782.					
KW	Protein identification; signal transduction pathway; metabolic pathway;					
KW	hybridisation assay; genetic mapping; gene expression control; promoter;					
KW	termination sequence.					
OS	Arabidopsis thaliana.					
PN	EP1033405-A2.					
XX	06-SEP-2000.					
PD	25-FEB-2000; 2000EP-0301439.					
XX	25-FEB-1999; 99US-0121825.					
PR	05-MAR-1999; 99US-0123180.					
PR	09-MAR-1999; 99US-0123548.					
PR	23-MAR-1999; 99US-0125788.					
PR	25-MAR-1999; 99US-0126264.					
PR	29-MAR-1999; 99US-0126785.					
PR	01-APR-1999; 99US-0127462.					
PR	06-APR-1999; 99US-0128234.					
PR	16-APR-1999; 99US-0128714.					
PR	19-APR-1999; 99US-0129845.					
PR	21-APR-1999; 99US-0130077.					
PR	23-APR-1999; 99US-0130449.					
PR	28-APR-1999; 99US-0130891.					
PR	30-APR-1999; 99US-0131449.					
PR	30-APR-1999; 99US-0132048.					
PR	04-MAY-1999; 99US-0132407.					
PR	05-MAY-1999; 99US-0132484.					
PR	06-MAY-1999; 99US-0132485.					
PR	06-MAY-1999; 99US-0132486.					
PR	07-MAY-1999; 99US-0132487.					
PR	11-MAY-1999; 99US-0132863.					
PR	14-MAY-1999; 99US-0134256.					
PR	23-JUL-1999; 99US-0134218.					
PR	23-JUL-1999; 99US-0145218.					

XX DHULST C, Ball S;
PI WPI; 2001-052291/07.
XX N-PSDB; AAC86955.
PT New recombinant nucleic acid encoding fusion of starch synthase and
PT second component, useful in pharmaceutical and food compositions, is
PT targeted to starch granules -
XX Claim 9; Fig 1; 52pp; French.
XX The present sequence represents a granule bound starch synthetase II
CC (GBSSI). The polynucleotide describes a recombinant nucleic acid,
CC comprising a polynucleotide encoding an adenosine diphosphate,
CC glucose-alpha1,4-glucan alpha4-glucosyltransferase or starch
CC synthetase. Placed upstream of a sequence that encodes a polypeptide
CC of interest. The adenosine diphosphate glucose-alpha1,4-glucan
CC alpha4-glucosyltransferase protein can migrate to sites of
CC biosynthesis of starch grains in plant cells, becoming associated
CC with these grains. The recombinant nucleic acid sequence is used to
CC target polypeptides of interest to starch grains.
XX Sequence 238 AA;
Query Match 32.8%; Score 1194; DB 22; Length 238;
Best Local Similarity 99.2%; Pred. No. 5.2e-86;
Matches 234; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 473 QLHAMHYGTPVAVASTGGLVDTVKEGVTGFHMGALNPKLDEADADALAAATVRRASEVFA 532
Db :|||||
3 RVHAMHYGTPVAVASTGGLVDTVKEGVTGFHMGALNPKLDEADADALAAATVRRASEVFA 62
QY 533 GGRYPEMVANCIQDLSWSKPAQKWEGLLEEVYVYGGCVATATKEETKVPVAEKIPGDLP 592
Db :|||||
63 GGRYPEMVANCIQDLSWSKPAQKWEGLLEEVYVYGGCVATATKEETKVPVAEKIPGDLP 122
QY 593 AVSYAPNTLKPVSASVEGNGAAAPKVGTTAPAMGAWRATTPSGPSAAATPKVITYKPAL 652
Db :|||||
123 AVSYAPNTLKPVSASVEGNGAAAPKVGTTAPAMGAWRATTPSGPSAAATPKVITYKPAL 182
QY 653 PATAKPTAGLKLAGEASTTSTSENGAASNGNGASAKTSAKPLVSAATKSA 708
Db :|||||
183 PATAKPTAGLKLAGEASTTSTSENGAASNGNGASAKTSAKPLVSAATKSA 238
RESULT 11
AAB49307
ID AAB49307 standard; Protein; 534 AA.
XX AAB49307;
XX 01-MAR-2001 (first entry)
XX Wheat starch synthase GBSS protein.
XX Wheat starch synthase; SSII; SSIII; starch content; starch synthesis;
KW food product; adhesive.
XX Triticum aestivum.
XX WO200066745-A1.
XX 09-NOV-2000.
XX 28-APR-2000; 2000WO-AU00385.
XX 29-APR-1999; 99AU-0000052.
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
PA (GOOD-) GOODMAN FIELDER LTD.
PA (LINA-) GRP LIMAGRAIN PACIFIC PTY LTD.
XX

PI MORELL M, Li Z, Rahman S, Appeals R;
XX WPI; 2000-647602/62.
XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and
PT WST-II, useful in modifying plant starch content and/or composition -
XX Example 15; Fig 9; 211pp; English.
XX The present invention relates to novel protein and coding sequences from
CC wheat. The proteins are wheat starch synthases, designated SSII and
CC SSIII. These can be used in the modification of plant starch content or
CC composition, and to screen plants to identify mutations which affect
CC starch content and composition. The starch can then be used in food
CC products, such as flour, and in films, coatings, adhesives, building
CC materials and packaging materials.
XX Sequence 534 AA;
Query Match 24.8%; Score 904; DB 21; Length 534;
Best Local Similarity 32.5%; Pred. No. 1.3e-62;
Matches 213; Conservative 48; Mismatches 103; Indels 292; Gaps 9;
QY 64 VAAEVAPWSKTGGGLDVTGGLPIELVKGHRVMTIAPDYQADAWDTVVYVIMGEKVR 123
Db :|||||
2 VGAEMAPWSKTGGGLDVTGGLPAPMAANGHRVMTIAPDYQADAWDTVVYVIMGEKVR 49
QY 124 YFHSIKKGVHRVWIDHPWFLAKVWKTGSKLYGPRSGADYLDNKRFAFCFAAEARV 183
Db :|||||
50 -----LEKVRGKTEKIYDPDAGTDYEDNOQRSLCQAALVPR 90
QY 184 L-----PF----- 186
Db :
91 LNLNPNFYFCHNHSYQGRSFDEFAQLNLPDRFKSSDFIDGYDPVEGRKISVYSEI 150
QY 187 -----GP-GEDCVFVANDWHSALVPVLLKDEYOPK 215
Db :|||||
151 KVVDKYERVYFHCYKRGVDRVFDHFCFSGPYGDEVVFCNDWHTGLLACYLKSQSN 210
QY 216 GQFTKAKSVLAHNIAFQGRWEEAFKDTLPQAAFDKLAFCSDYAKVYTEATWEDEK 275
Db :
211 GIYRAAKVA----- 219
QY 276 PPLTGTYKKINWLKGGIIAADKLVTSPNYATEIAAAGGVELDTVIRAKGIEGVNG 335
Db :|||||
220 -----NWKAGILQADKVLTVSPYYAEELISGEARGCELDNMLRTGLTITVNG 268
QY 336 MDIEENPKTKDFLSAPYDQNSVYAGAAK----- 366
Db :|||||
269 MDVSEWDPDKDFLAVNYDITTALEKALNKEALEKALNKLKEDVQIVLLGTGKKKE 328
QY 367 -----EALQAEGLPVDPTAPLPAFAGRLLEEOKGVLDILA 401
Db :|||||
329 RLLKSTEEKPSKVRVAVRNAPLAELQAEVGLVDKRVPLVAFVIGRLEEOKGVLDILA 388
QY 402 ALPKLILATPKVQIALLGTGKAAYEKLVAIGTKYKRAKGVKWFSAPLHMLTAGADFML 461
Db :
389 SIPEI-----HQMAGADVLA 404
QY 462 VPSRFEPCGLIQIHAMHYGTPVAVASTGGLVDTVKEGVTGFHMGALNPD---KLDEADADA 519
Db :|||||
405 VTSRFEPCGLIQIQQMRGTGTPCAGSTGGLVDTIVEGKTGFHMGRLSYDCNVVPEADVK 464
QY 520 LAATVRRASEVFAGGRYPE-----MVANCISODLSWSKPAQKWEGLLEEV 564
Db :|||||
465 VVTLKRAVKVGTGPAYHEAPLAVENVAAPMVKNCMIODLSWKGPAKNWEDVLEL 520
RESULT 12
AAY09004
ID AAY09004 standard; Protein; 647 AA.
XX AAY09004;
AC AAY09004;

XX 05-JUL-1999 (first entry)
 XX Wheat starch soluble synthase I (SSS I) amino acid sequence.
 DE
 XX starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; CBSS;
 KW starch branching enzyme; starch soluble synthase; debranching enzyme;
 KW endosperm; wheat; barley; granule-bound synthase; glutenin; starch;
 KW grain softness protein I; bacterial isoamylase; glycogen synthase;
 KW WSBE I-D4 gene.
 XX
 XX Triticum tauschii.
 OS
 XX WO9914314-A1.
 PN
 XX 25-MAR-1999.
 PD
 XX 11-SEP-1998; 98WO-A000743.
 PF
 XX 20-MAR-1998; 98AU-0002509.
 PR
 XX 12-SEP-1997; 97AU-0009108.
 XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PA (GOOD-) GOODMAN FIELDER LTD.
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
 PA (AUSU) UNIV AUSTRALIAN NAT.
 XX
 PI Li Z, Morell M, Rahman S;
 XX
 XX WPI; 1999-229525/19.
 DR
 XX N-PSDB; AAX34651, AAX34652.
 DR
 XX New isolated cereal plant enzyme genes used for, e.g. expression of
 PT antisense sequences of granule bound synthase
 PT
 XX
 PS Claim 13; Page 95-97; 171pp; English.
 XX
 CC The invention relates to a novel enzyme of starch biosynthetic pathway
 CC in a cereal plant, where the enzyme is selected from starch branching
 CC enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching
 CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or
 CC SBE I of rice or maize. The methods and products can be used for
 CC targeting expression specifically to the endosperm of the seeds of cereal
 CC plants such as wheat or barley. They can be used for the expression of
 CC e.g. antisense sequences of granule-bound synthase (CBSS), SBE II, low
 CC mol. wt. glutenin, grain softness protein I, bacterial isoamylase,
 CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They
 CC can be used for modifying the characteristics of starch produced by a
 CC plant. The present sequence represents the wheat SSS I protein sequence.
 CC
 XX Sequence 647 AA;
 XX
 Query Match 24.7%; Score 900.5; DB 20; Length 647;
 Best Local Similarity 37.8%; Pred. No. 3.2e-62;
 Matches 225; Conservative 94; Mismatches 205; Indels 71; Gaps 18;
 QY 4 ASTSRPSSARPIVINAASFGVKITANOLLRELARGSARKSTSAVTGATGATCALD--- 60
 DB 78 APAGSPATOPPLPDA---GVGELAPDLLE---GIAEDSIDSIIVAASEQDSIMDANE 131
 QY 61 -----IVMAAEVAPWSKGTGLGDTVGGGLPIELVKGRHVTMTAPRY-----DQY 105
 DB 132 QPOAKVTRSFVFTGEAPYAKSGGLGDTGCGSLPIALAARGHVRVVMVMPRYLNGSSDKNY 191
 QY 106 ADAMDTSVVDIM---GEKRYEPHSIKKGVRHWIDHPWFLAKWKGKTSKLYGPRSGA 161
 DB 192 AKALYTGKHKIKPCFGSGSHVTFPEHYRDNDVWVFDHP-----SYHRPGS-LYGDNFGA 245
 QY 162 DYLDNHRKREAFCKAAEAAKVLDPGP---GEDCVFVANDWHSALVPVLLKDEYQPGQF 218
 DB 246 -FGDNQFRYTLCLYAAACEAPLILEGGYIYGONCMFVYNDWHSALVPVLLAAKRYPGYV 304
 QY 219 TKAKSVLAHNIAPQGRWWEAEAFDKTLPOAFDKLAFSDGYAKVYITEATPMEDEKPL 278

DB 305 RDSRSTLVJHNLAHQGLEPASTYPDLGLPPEWYGALEW-----VPEWARRHALDKG-- 356
 QY 279 TGTYYKKINLKGIIAADKLVTSNPNYATELAADAAGGVELDTVI--RAKGIEGIVNGM 336
 DB 357 -----EAVNFLKAVTADRIVTVSOGYSWEVIT-AEGGGLNELLSRKSVLNGIVNGI 410
 QY 337 DIEWNPKTDKFLSAPYDONSIVYAGKAAAKALQAEIPLGPDPTADLPAFIFGRLEQKGY 396
 DB 411 DINDWNPETDKCLPHHYSVDDI-SGAKCKAEIQLKELGFLFVREDVPLIGFIRGLDYOKGI 469
 QY 397 DIILAALPKLITPKVQIAIILGTGKAAEYKLVNAIGTKYKGRAGVVKFSAPLAHMLTAG 456
 DB 470 DLIKMAIPELM-REDQFVNLGSGDPIFEGWMHSTESSKDKFRGWVGSFVPSVSHRITAG 528
 QY 457 ADFMLVPSRFEPCGLIQLHAMHYGTVPVVASTGGLVDTVK-----EGVTGFFHMGAL 507
 DB 529 CDILLMPSRFEPCGLNQLYAMQYGTVPVVGHTGGLDRTVETENPFPGAKGEESTGWAFSPL 588
 QY 508 NPKLDEADADALAAIVRRASEVFAGRYPEAVANCISQDLSSKSPAKQWEGGLE 562
 DB 589 TVDKMLWALRTAM-STFREHKPSWEG-----LMKRGMTKDHTWDHAAEQYEQIFE 637

RESULT 13

AB892160
 ID AB892160 standard; Protein; 792 AA.

XX AC AB892160;

XX 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 1371.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP09892.

XX 28-AUG-2001; 2001WO-EP09892.

XX (FARB) BAYER AG.

XX Tietjen K, Weidler M;

XX WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -

XX Claim 5; SEQ ID NO 1371; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins
 CC (AB890790-AB894016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.

XX Sequence 792 AA;

Query Match 24.7%; Score 898; DB 23; Length 792;
 Best Local Similarity 40.5%; Pred. No. 6.9e-62;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:47:01 ; Search time 18.3001 Seconds
(without alignments)
3719.278 Million cell updates/sec

Title: US-09-980-771A-3
Perfect score: 3641
Sequence: 1 MAVASTSRPSSARPVINA.....SASKTSAKPLVSAATRKSA 708

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1485.5	40.8	608	2 S43341	starch synthase (E
2	1463.5	40.2	607	1 YUPOY	starch synthase (E
3	1431.5	39.3	603	2 S61504	glycogen(starch) s
4	1427	39.2	609	2 J00703	glycogen(starch) s
5	1426	39.2	609	2 S11481	glycogen(starch) s
6	1421.5	39.0	603	1 YUBHY	glycogen(starch) s
7	1417	38.9	608	2 T10906	starch synthase (E
8	1405.5	38.6	610	2 T86453	granule-bound star
9	1376	37.8	608	2 T14731	glycogen(starch) s
10	1372	37.7	615	1 YUWTY	glycogen(starch) s
11	1369.5	37.6	605	1 S07314	glycogen(starch) s
12	1194	32.8	238	2 T07921	glycogen(starch) s
13	903	24.8	641	2 T07668	probable starch sy
14	874	24.0	752	2 S61505	starch synthase (E
15	873.5	24.0	626	2 J02322	glycogen(starch) s
16	873	24.0	732	2 T01208	starch synthase (E
17	869.5	23.9	622	2 T01414	starch synthase (E
18	869	23.9	610	2 T06280	starch synthase (E
19	849.5	23.3	788	2 T07667	probable starch sy
20	847.5	23.3	698	2 T01209	glycogen(starch) s
21	839	23.0	491	2 T06798	starch synthase (E
22	684.5	18.8	477	2 T95130	probable starch sy
23	683.5	18.8	477	2 H98000	glycogen synthase
24	677.5	18.6	484	2 S40051	starch synthase (E
25	669	18.4	483	2 C86712	starch synthase (E
26	652	17.9	480	2 H98228	starch synthase (E
27	652	17.9	480	2 AF3057	glycogen synthase
28	651	17.9	486	2 H72321	glycogen synthase
29	637	17.5	477	2 A97176	glycogen synthase,

30	622	17.1	484	2 F82165	glycogen synthase
31	614.5	16.9	476	2 E83785	starch (bacterial
32	604.5	16.6	486	2 C96018	probable starch sy
33	604	16.6	476	2 AE0479	starch synthase (E
34	590.5	16.2	530	2 AH3194	glycogen synthase
35	587	16.1	477	1 SYEGL	starch synthase (E
36	587	16.1	477	2 B98163	glycogen synthase
37	587	16.1	477	2 C86009	glycogen synthase
38	587	16.1	477	2 AI0995	glycogen synthase
39	570.5	15.7	1230	2 T07663	starch synthase (E
40	569	15.6	476	2 C64119	soluble starch syn
41	564.5	15.5	1025	2 H86250	starch synthase (E
42	560.5	15.4	472	2 AI2040	hypothetical prote
43	552.5	15.2	463	2 C70363	glycogen synthase
44	536	14.7	492	2 AG1810	glycogen (starch)
45	535	14.7	477	2 S76496	hypothetical prote

ALIGNMENTS

RESULT 1
S43341
starch synthase (EC 2.4.1.21) precursor - cassava
N:Alternate names: bacterial-glycogen synthase; glycogen synthase; starch synthase
C:Species: Manihot esculenta (cassava)
C:Date: 07-Sep-1994 #sequence_revision 03-Nov-1995 #text_change 16-Aug-2002
C:Accession: S43341
R:Salehuzzaman, S.N.I.M.; Jacobsen, E.; Visser, R.G.F.
Plant Mol. Biol. 23, 947-962, 1993
A:Title: Isolation and characterization of a cDNA encoding granule-bound starch synth
A:Reference number: S43341; MUID:94083565; PMID:8260633
A:Accession: S43341
A:Molecule type: mRNA
A:Residues: 1-608 <SAL>
A:Cross-references: EMBL:X74160; NID:g437041; PIDN:CAA52273.1; PID:g437042
C:Genetics:
A:Gene: GRSS; waxy
A:Genome: nuclear
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose produc
C:Superfamily: starch synthase
C:Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltra
F:1-78/Domain: transit peptide (amyloplast) #status predicted <TNP>
F:79-608/Product: ADPglucose-starch glucosyltransferase #status predicted <MAT>

Query Match	40.8%	Score	1485.5	DB	2	Length	608
Best Local Similarity	53.6%	Pred. No.	1.1e-80	Mismatches	67	Indels	39
Matches	304	Conservative	157	Gaps	11		
QY	42	KSTSRSAVTG----	ATGATC--	ALDIVMVAEEVAPWSKTTGGDVTGGTGLPIELVKRGHRV	95		
Db	59	KAVKVSATGNGRPAKILICGHGMNLI	FVGAEGVPSKTTGGDVLGLPPMAARGHRV	118			
QY	96	MTIAPRYQYADAMDTSVVVDI-MG----	EKVYFHSIKKGVHRVWIDHPWELAKVWGKTG	151			
Db	119	MTVSPRYQYDAMDTSVSVEIKIGDRIET	VRFFHSYKRGVDRVFDHPMLEKVGKTG	178			
QY	152	SKLVGRSGADYLDNHRKFALECKAIEAARVLPF-----	GP-GEQCVFVANDWHSAL	203			
Db	179	SKIVGPRAGLDYQDNQLRFLSLCLLAALPAKVLNLSKNSKFSFGPYGEEVAFIANDWHTAL	238				
QY	204	VPVLLKDEYQKPGQFTAKSVLAHNI	AFQGRMWEAEFKTKLPQAAFDKLAESDGYAKV	263			
Db	239	LPVLYKALYQPMGIYKHKAVAFCH	NIAYQGRFAFSDPRLNLPDKFKSSDFIDGDK-	297			
QY	264	YTEATPMEEDKPPLTGKTGYKKINW	LKGLIAADKLTVSPNYATEIAADAAGGVGLDTV	323			
Db	298	-----PVKG----	RKINMKAGILESRLVTSPPYAOEIVSGVERGVGLDNF	341			
QY	324	IRAKGIEGVNGMDIEFNPNKTKFLS	APYDQNSVYAGAKAAKALQAEGLPVDPTAPL	383			
Db	342	IRKTGIAGIINGMDVQVWNPVTDKYID	IHYDATTVMDAKPLKALQAEVGLPVDNRVPL	401			

```

384 FARTGRLEOKGVDIILAAALPKILATPKVOIALILGTGKAAYEKLVNALIGTKYKGRAGVY 443
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 402 IGFTRLEOKGSDIFVAAALSQ-LVHNQVILILGTGKKFKEKOEHLLEVLYDPKARGVA 460
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 444 KFSAPLAHMITAGADFMVPSRFPCCGLIQLHAMHYGTVPVASTGGSLDVTVKEGVTGFH 503
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 461 KFNYPPLAHMITAGADFMVPSRFPCCGLIQLHAMRYGTVPVASTGGSLDVTVKEGYTFQ 520
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 504 MGALNP--DKLDEADALAAVTRRASEVFAGGRYPDEMVA NCISODLSWSKPAOKWEGLL 561
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 521 MGALHVECDKIDSADVAIVKTVARALGTATYATALREMLNCMAQDLSWKGPARMWEKML 580
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 562 EEVVYGVKGAVATAKKEIKVPVAKIP 588
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 581 LDLEV-TGSEPTGEGEELAPLAKENVP 606
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
YUPOY
starch synthase (EC 2.4.1.21) precursor - potato
A:Alternate names: starch synthase
C:Species: Solanum tuberosum (potato)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002
C:Accession: S16555; S24392; S26060; S26061
R:van der Leij, F.R.; Visser, R.G.F.; Ponstein, A.S.; Jacobsen, E.; Feenstra, W.J.
Mol. Gen. Genet. 228, 240-248, 1991
A:Title: Sequence of the structural gene for granule-bound starch synthase of potato (Solanum tuberosum)
A:Reference number: S16555; MUID:91360072; PMID:1886609
A:Accession: S16555
A:Molecule type: DNA
A:Residues: 1-607 <LEI>
A:Cross-references: EMBL:X58453; NID:g21470; PIDN:CAA41359.1; PID:g21471
A:Note: the authors translated the codon AAC for residue 453 as Gly and GCT for residue 454
A:Accession: S24392
A:Molecule type: protein
A:Residues: 78-92, 'X', '94-98, 'XXX', 102, 'XX', 105-107 <LE2>
R:Roche, W.; Becker, D.; Kull, B.; Salamini, F.
J. Genet. Breed. 44, 311-315, 1990
A:Title: Structural and functional analysis of two waxy gene promoters from potato.
A:Reference number: S26060
A:Accession: S26060
A:Molecule type: DNA
A:Residues: 1-43 <ROH1>
A:Cross-references: EMBL:X52416; NID:g21613; PIDN:CAA36667.1; PID:g21614
A:Experimental source: cv. Granola, clone G1
A:Accession: S26061
A:Molecule type: DNA
A:Residues: 1-43 <ROH2>
A:Cross-references: EMBL:X52417; NID:g21615; PIDN:CAA36668.1; PID:g21616
A:Experimental source: cv. Granola, clone G28
C:Genetics:
A:Gene: waxy
A:Genome: nuclear
A:Introns: 111/3; 138/3; 171/3; 201/3; 223/1; 256/3; 293/2; 374/3; 433/3; 497/3; 526/3;
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producing a
A:Pathway: starch biosynthesis
C:Superfamily: starch synthase
C:Keywords: amyloplast; glycogen/starch biosynthesis; glucosyltransferase; hexosyltransferase
F:1-77/Domain: transit peptide (amyloplast) #status predicted <INF>
F:78-607/Product: ADPglucose-starch glucosyltransferase #status experimental <MAT>

Query Match 40.2%; Score 1463.5; DB 1; Length 607;
Best Local Similarity 52.7%; Pred. No. 2.2e-79;
Matches 298; Conservative 70; Mismatches 157; Indels 41; Gaps 10;

QY 39 SARKSTSRSAVTCATGATCALDIVMVAEVAFWKSTGGGLDVTGGLPIELVKGHRVMTI 98
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ASRTETKRPCSATIVCGKGMNLIIFVTEVPWPSKTGGGLDVLGGLPALAARGHRVMTI 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 99 AFRYDQYADWNTSVVVDI-MG---EKVRFHSHKGVHRVWIDHPWFLAKYWGKTSKL 154
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 121 SPRYDQYKDWNTSVAVEVKVGDSTIEIVRFHCHYKRGVDRVDFVDPHMFLEKVMWKGTSKI 180
QY 155 YGPRSGADYLDNHRKRALFCKAAEAARVLPF-----GP-GEDCVFVANDHMSALVPV 206
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 YGPRAGLDYLDNELRFSLLCOALEAPKVLNNSNYSPYGYGVDFLFIANDWHTALIPC 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 207 LLKDEYQPKGQFTKAKSVLAITHNIAFGQRMWEEAFKOTKLPQAAFDKLAFSQGYAKVYTE 266
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 YLKSMYQSRGYLNAKVAFCIHNIAYGQGRFSFSDPPLNLNLPDEFGRGSDFDIDGYEK---- 296
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 267 ATPMEEDKPLTGTGYKKINWLGIIIAADKLVTSVSNYATEIAADAAGGVVELDTVIRA 326
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 297 -----PVKG---RKINMKAGIILSHRVTVSPYIAGELIVSAVDKGVGLDSVLRLK 343
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 327 KGIEGIVNGMDIEENPKTDKFLSAPYDQNSYIAGKAAKALQALQELGLPVDPTPLPAF 386
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 344 TCITGIVNGMDTOENPATDKYDVTYKIDITVMDAKPLLKEALQAAVGLPVDKKTPLIGF 403
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 387 IGRLEEQKGVDIILAAALPKILATPKVOIALILGTGKAAYEKLVNALIGTKYKGRAGVVKFS 446
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 404 IGRLEEQKGSIDLVAALHFKTGL-DVQIVVLGTGKKEFEQETEQLLEVLYDPNKAAGVAKFN 462
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 447 APLAHMITAGADFMVPSRFPCCGLIQLHAMHYGTVPVASTGGSLDVTVKEGVTGFHMA 506
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 463 VPLAHMITAGADFMVPSRFPCCGLIQLHAMRYGIVPICASTGGGLDVTVKEGYTFGHMA 522
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 507 LNP--DKLDEADALAAVTRRASEVFAGGRYPDEMVA NCISODLSWSKPAOKWEGLL--- 561
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 523 FNVECDVVDPADVLKTVITVARALAVYGTFLAEAIKNCMSPELSKPEPAKKWETLILGL 582
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 562 -----EEVVYGVKGAVATAKKEIKVP 582
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 583 GASGSEPGVGEF-EIAPLAKENVATP 607
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
S61504
glycogen(starch) synthase (EC 2.4.1.11) isoform I precursor - garden pea
N:Alternate names: glycogen(starch) synthase
C:Species: Pisum sativum (garden pea)
C:Date: 20-Jul-1996 #sequence_revision 07-Feb-1997 #text_change 16-Aug-2002
C:Accession: S61504; S72372
R:Dry, I.; Smith, A.; Edwards, A.; Bhattacharya, M.; Dunn, P.; Martin, C.
Plant J. 2, 193-202, 1992
A:Title: Characterization of cDNAs encoding two isoforms of granule-bound starch synt
A:Reference number: S61504; MUID:93251108; PM.D:1302049
A:Accession: S61504
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-603 <DRY>
A:Cross-references: EMBL:X88789; NID:g887570; PIDN:CAA61268.1; PID:g887571
A:Accession: S72372
A:Molecule type: protein
A:Residues: 76-77, 'X', 79, 81-88 <DRW>
C:Superfamily: starch synthase
C:Keywords: glycosyltransferase; hexosyltransferase
F:1-75/Domain: signal sequence #status predicted <SIG>
F:76-603/Product: glycogen (starch) synthase isoform I #status experimental <MAT>

Query Match 39.3%; Score 1431.5; DB 2; Length 603;
Best Local Similarity 50.1%; Pred. No. 1.8e-77;
Matches 291; Conservative 85; Mismatches 158; Indels 47; Gaps 13;

QY 26 KTAQQLRLRELARGSARKSTSRSAVTCATGATCALDIVMVAEVAFWKSTGGGLDVTGGLP 85
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 46 RSLNKLHVTRATARGSSDTSEKSLGK--IVCGMSLVFVGAEVGPMWPKTGGGLDVLGGLP 103
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 86 IELVKGHRVMTIAPRYDQYADWNTSVVVDI-MGEK---VRYFHSITKGVHRVWIDHPW 141
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 104 PVLAGNHRVMTIAPRYDQYADWNTSVVVDI-MGEK---VRYFHSITKGVHRVWIDHPW 163
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 142 FLAKVWGKTSKLYGPRSGADYLDNHRKRALFCKAAEAARVLPF-----GP-GEDCV 193
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```


Db 164 FLERVWGKTGSKLYGPKTGIDYRDNLRLRFSLICQAALEAPVNLNLSKSYFSGPYGEDVI 223
 QY 194 EVANDHSAVLPVLLKDEYQPKQFTYKAKSVLAHNTAFQGRMWEAFKD---TKLPQAA 250
 Db 224 FVANDHSAVLPVLLKDEYQPKQFTYKAKSVLAHNTAFQGRMWEAFKD---TKLPQAA 250
 QY 251 FDLAFSDGAKYVTEATPMEDEKPLUQTKKINLWKGIIAADKLVTSPFNATEI 310
 Db 281 RSSFDIDGYNK-----PCGE---KKINWKGIIAADKLVTSPFNATEI 310
 QY 311 ADAAGVELDVTIRAKGIEGVNMDIEENPKTDKFLSAPYDQNSVYAKAAKALQ 370
 Db 324 ISGEDRGVELDNIIRSGIIGVNGMDNREWSPTDRIYIDVHYNETTVEAKPLKGLQ 383
 QY 371 AELGLPDPAPLFAFAGRLLEQKGYDIIAALPKILATPKVQIATILGTCAAYEKLVNA 430
 Db 384 AELGLPDPAPLFAFAGRLLEQKGYDIIAALPKILATPKVQIATILGTCAAYEKLVNA 430
 QY 431 IGTYKGRAGGVVFKFAPLAHMLTAGDFMLVPSRFEPCGLIOLHAMHYGTVPVASTGG 490
 Db 443 LEKYPKAIGTKFNSPLAKHIIAGADFIVSRFEPCGLIOLHAMHYGTVPVASTGG 490
 QY 491 LVDTVKGVTFMGALNP--DKLDEADADALATVRRASEVFAGGRYPFWVANCISQDL 548
 Db 503 LVDTVKGVTFMGALNP--DKLDEADADALATVRRASEVFAGGRYPFWVANCISQDL 548
 QY 549 SWSKPAQKVEGLEEV-----VYKGG--VATAKKEEIKVP 582
 Db 563 SWSKPAQKVEGLEEV-----VYKGG--VATAKKEEIKVP 582

RESULT 4

JQ0703
 N:glycogen(starch) synthase (EC 2.4.1.11) - rice
 C:Species: Oryza sativa (rice)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Aug-2002
 C:Accession: JQ0703
 R:Wang, Z.Y.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.
 Nucleic Acids Res. 18, 5898, 1990
 A:Title: Nucleotide sequence of rice waxy gene.
 A:Reference number: JQ0703; MUID:91016948; PMID:2216792
 A:Accession: JQ0703
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-609 <WAX>
 A:Cross-references: EMBL:X53694
 A:Experimental source: strain subsp. japonica Hengfeng
 C:Genetics:
 A:Gene: waxy
 A:Introns: 114/1; 141/1; 174/1; 204/1; 225/2; 259/1; 295/3; 377/1; 436/1; 500/1; 529/1;
 A:Function:
 A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
 C:Superfamily: starch synthase
 C:Keywords: chloroplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltrans

Query Match 39.2%; Score 1427; DB 2; Length 609;
 Best Local Similarity 48.8%; Pred. No. 3.3e-77;
 Matches 301; Conservative 82; Mismatches 172; Indels 62; Gaps 13;
 QY 3 VASTSRPSS-----ARPIVINAASFVGVKKTAN---QLRLARGSAKSTSRSA 48
 Db 18 IADRSAPSLLRHFGQGLKPRSPAGDATSLSVTSARATPKQQRSVQKRSR---FPVS 74
 QY 49 VNGATGATCALDIVMVAEVAEPWSTGGGLDVTGGLPIELVKRGRHVMITAPRYQYADA 108
 Db 75 VVIATCA--GMNVFVGAEMAPWSTGGGLDVTGGLPLPAPMAAANGHRVMVSPRYQYKDA 132
 QY 109 WDTSSVVVDI---MGEKRYFHSIKGVHVRWIDHPWFLAKYWGKTSKLYGPRGADYL 164
 Db 133 WDTSSVVAEVKADRYERVRFFHCYKRGVRDVFIDHPSFLEKYMVGKTGKIYGPDTGVYK 192
 QY 165 DNHRKRFALFCKAAIEAARVL-----PFGP---GEDCVFVANDHSAVLPVLLKDEYQPKG 216

Db 193 DNQRFSLICQAALEAPVNLNLSKSYFSGPYGEDVI 223
 QY 217 QFTKAKSVLAHNTAFQGRMWEAFKD---TKLPQAA 250
 Db 253 IYRNKAVAFCHNLSYQGRFAFEDYPELNLSEFSSFDIDG-----DTPVEG--- 302
 QY 277 PLTGKTYKKINLWKGIIAADKLVTSPFNATEIADAAGVELDVTIRAKGIEGVNMG 336
 Db 303 -----RKINWKGIIAADKLVTSPFYAAEELISGLARGCELDNMLRTGITGVNMG 355
 QY 337 DIEENPKTDKFLSAPYDQNSVYAKAAKALQELGLPDPAPLFAFAGRLLEQKGY 396
 Db 356 DVSMDPSKDIITAKYDAITAEKALNEALQAEAGLPVDRKIPLTAFAGRLLEQKGP 415
 QY 397 DIILAALPKILATPKVQIATILGTCAAYEKLVNAIGTYKGRAGGVVFKFAPLAHMLTAG 456
 Db 416 DVMAAATPELM--QEDVOIVLLGTGKKFELKLSWEEKYFGKRVAVVKNAPLAHLIMAG 474
 QY 457 ADMLVPSRFEPCGLIOLHAMHYGTVPVASTGGIOLHAMHYGTGFHMGALNPD--KLDE 514
 Db 475 ADLVAPSRFEPCGLIOLGMRGTACACASTGGIOLHAMHYGTGFHMGALNPD--KLDE 514
 QY 515 ADADALATVRRASEVFAGGRYPFWVANCISQDL SWSKPAQKVEGLEEVVYKGG---- 570
 Db 535 SDVKVAAATILKRAIKVVGTPAYEEMVRNQMQLSWKGPAPKNWENVL--LGLGVAGSAPG 592
 QY 571 -----VATAKKEEIKVP 582
 Db 593 IEGDETAPLAKENVAAP 609

RESULT 5

S11481
 N:glycogen(starch) synthase (EC 2.4.1.11) precursor - rice
 C:Species: Oryza sativa (rice)
 C:Date: 05-Mar-1995 #sequence_revision 10-Nov-1995 #text_change 16-Aug-2002
 R:Wang, Z.Y.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.
 Nucleic Acids Res. 18, 5898, 1990
 A:Title: Nucleotide sequence of rice waxy gene.
 A:Reference number: JQ0703; MUID:91016948; PMID:2216792
 A:Accession: S11481
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-609 <WAX>
 A:Cross-references: EMBL:X53694; NID:g577598; PIDN:CAA37732.1; PID:g577599
 R:Okagaki, K.J.
 Plant Mol. Biol. 19, 513-516, 1992
 A:Title: Nucleotide sequence of a long cDNA from the rice waxy gene.
 A:Reference number: S22519; MUID:92322986; PMID:1377969
 A:Accession: S22519
 A:Molecule type: mRNA
 A:Residues: 1-609 <OKA>
 A:Cross-references: EMBL:X62134; NID:g20402; PIDN:CAA44065.1; PID:g20403
 R:Hirano, H.Y.; Sano, Y.
 Plant Cell Physiol. 32, 989-997, 1991
 A:Title: Molecular characterization of the waxy locus of rice (Oryza sativa).
 A:Reference number: S30485
 A:Accession: S30485
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-609 <HIR>
 A:Cross-references: EMBL:X58228; NID:g20400; PIDN:CAA41186.1; PID:g20401
 A:Accession: PC2190
 A:Molecule type: protein
 A:Residues: 78-113 <HIR>
 A:Experimental source: leaf, cDNA pOSLHC2120
 C:Comment: This protein is involved in amylose synthesis in the rice endosperm.
 C:Genetics:
 A:Gene: waxy
 A:Introns: 113/3; 140/3; 173/3; 203/3; 225/1; 258/3; 295/2; 376/3; 435/3; 499/3; 528/

C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
A:Pathway: starch biosynthesis
C:Superfamily: starch synthase
C:Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glycosyltransferase;
F:1-77/Domain: transit peptide (amyloplast) #status predicted <TNP>
F:78-609/Product: glycogen (starch) synthase #status experimental <NAT>

Query Match	39.2%;	Score 1426;	DB 2;	Length 609;
Best Local Similarity	48.9%;	Pred. No. 3.8e-77;		
Matches	302;	Conservative 172;	Indels 62;	Gaps 13;
QY	3	VASTSRSS-----ARPIVINAAGPGVKKTAN--	OLDRELARGSARKSTSRSA	48
DB	18	IADRSAPSSLLRHGFOGLKPRSPAGGDATSLSVTTSARATPKQORSVQGRSR---	PPSV	74
QY	49	VTGATGATCALDIYVMAAEVAPSKTGGGLGVDVTGGLPIELVKRGHRVMTIAPRDQYADA	108	
DB	75	VVYATGA--GMNVYFVGAEMAPSKTGGGLGVDLGLPLPAMAANGHRVYVISPRIYDQYKDA	132	
QY	109	WDTSSVVYDI----NGEKVRYPHSKKGVRHWIDHPWFLAKVCKTKGSKLGPGRSGADYL	164	
DB	133	WDTSVVAEIKVADRYERVRFPHCYKRGVDYRFDIDHPSFELEKVGWKTGKTIYGPDTGVYDK	192	
QY	165	DNHKRFALFCKAAIEBARVL-----PFGP-----GEDCVFVANDHMSALVPVLLKDEYQPKG	216	
DB	193	DNQMRFSLLCQAALEAPRILLNNPYPFGTYGEDVVFVVCVNDWHTGPLASYLKNNYQPNG	252	
QY	217	QFTAKSVLAITHNTAFQGRWEEFAFKDKLPQAAFDKLAFSDGAKVYVTEATPMDEDEKP	276	
DB	253	IYRNAKVAFCIHNLSIQGRFAFEDYPELNLSEFRSSDFDIDGY-----DTPVEG----	302	
QY	277	PLTGKTYKINWLKGGIIAADKLVTVSPNYATEIAADAAGVGVELDTVIRAKGIEGVNMG	336	
DB	303	-----RKINWMKAGILEADRVLTVSPYYAEELISGIARGCELDNIMRLGITGIVNMG	355	
QY	337	DIEWNPKTKDFLSPYDQNSYAGKAAKAEALQAEGLPVDPTAPLFAFTGRLEEOKGV	396	
DB	356	DVSEWDPKDKYITIAKYDATTATBAKALNKEALQAEAGLPVDRKTIPLFAFTGRLEEOKGP	415	
QY	397	DIIILAAPLKILATPKVQITAILGTGKAAAEKVLVNAITGTYKGRAGKGVWFSAPLAHMLTAG	456	
DB	416	DVMAAAIPELM--QEDVQIVLLGTGCKKFEKLLKSMEEKYPCKVRVAVVKNAPLAHLINAG	474	
QY	457	ADEMLVPSRFPFCGLIQLHAMHYGHVPVVAVTGGLVDTVKCEGVTGFHMGALNPD--KLDE	514	
DB	475	ADVLAVPSRFPFCGLIQLQGMRYGTGTPCASCATGGGLVDTVIECKTGTFHMGRLSVDCKVYEP	534	
QY	515	ADADALATVRASVFAAGGYPPEMVANCIQSOLISWSKPAOKWEGLEEVYVYKGG-----	570	
DB	535	SDVKVAATLAKRAIKVGTGPAYEEMVRNMQNDLLSWKGPKNWENVL--LGLGVAGSAPG	592	
QY	571	-----VATAKKEEIKVP	582	
DB	593	LEGDEIATPLAKENVAAP	609	

```

RESULT 6
YUHHY
glycogen(starch) synthase (EC 2.4.1.11) precursor - barley
N:Alternate names: starch synthase
C:Species: Hordeum vulgare (barley)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002
C:Accession: S01727; S01728
R:Rohe, W.; Becker, D.; Salamini, F.
Nucleic Acids Res. 16, 7185-7186, 1988
A:Title: Structural analysis of the waxy locus from Hordeum vulgare.
A:Reference number: S01727; PMID:88403345; PMID:2970062
A:Accession: S01727
A:Molecule type: DNA
A:Residues: 1-603 <ROH>
A:Cross-references: EMBL:X07931; NID:g19126; PIDN:CAA30755.1; PID:g295809
A:Accession: S01728

```

A; Status: translation not shown
 A; Molecule type: mRNA
 A; Residues: 1-603 <ROH2>
 A; Cross-references: EMBL:X07932; NID:g19128; PIDN:CAA30756.1; PID:g19129
 C; Genetics:
 A; Gene: waxy
 A; Introns: 106/3; 133/3; 166/3; 218/1; 251/3; 369/3; 429/3; 493/3; 522/3; 565/3
 C; Function:
 A; Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
 A; Pathway: starch biosynthesis
 C; Superfamily: starch synthase
 C; Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltra
 E: 1-72/Domain: transit peptide (amyloplast) #status predicted <INP>
 E: 71-603/3product: UDPglucose-starch glycosyltransferase #status predicted <MAT>

Query Match	39.04;	Score 1421.5;	DB 1;	Length 603;
Best Local Similarity	51.08;	Pred. No. 7e-77;		
Matches	291;	Conservative 76;	Mismatches 167;	Indels 37; Gaps 10;
Qy	8	RPSSARPVINAASFGVKKTANOLLRELARGSARKSTSRSVATCATGATCATCALDIVMVAEE	67	
Db	31	RPRPADALCMRTIGASAAPKQ-SRAHGRSRR---CLSVVVSATCS---GMNLVYFGAE	84	
Qy	68	VAPSKGTGGLDVTGGPLIELVKGHRVMTIAPRYDQYADAWDTSVVVDIM----GSKVR	123	
Db	85	MAPWSKTGGLDVLGGLPPAMAANGHRVMVSPRYDQYKDAWDTSVISEIKVADEYERVR	144	
Qy	124	YFHSIKCKVHRVWDHDPWFLAKVWGKTSKLYGPRSGADYLDNHKRALPCKAAIEARV	183	
Db	145	FFHCYKRGVDKVFIDHPWFLEBKVRGKTEKIIYGDGADHYEDNQQRSLLCOAALAPRI	204	
Qy	184	L-----PR--GP-GEBCYFVANDMHSALVPVLKDEYQPKQGFKAKSVAIHNIATQGR	235	
Db	205	LNLNNPYESPGYGEDVVFVCDNWHGTGLLACVLSKNSYQSGIYRTAKVAFCHINLSYQGR	264	
Qy	236	MWEAFKDTKLPOAFDKLAFSDGKAVKYVTEATPWEDEKPLPKYKTKYKKNLWLGKGLIA	295	
Db	265	FSFDDFAQLNLPDRKFSFDFIDGYDK-----PVGE---RKINMKAGTILQ	307	
Qy	296	ADKLVTSPNVAETEIAADAAGVELDTVIRAKGTGIVNGMDIEWNPKTDKFLSAPYDQ	355	
Db	308	ADKLVTSPYYAAEELISGEARGCELDNTRLTGTGTGIVNGMDVSEWDTKDKFLAVNYDI	367	
Qy	356	NSVIAGKAAAEALQAEGLGDPDTPAPLAFIGRLEEQKGVDIILAALPKILATPKVOIA	415	
Db	368	TTALEKALKNEALQAEVGLPDRKVPVLAFIGRLEEQKGPDMVIAAIPETILKEEDVGII	427	
Qy	416	ILGTGKAAYEKLIVNAIGTKYGRAGVYKFSNAPLAHMLTAGADFWLPSRREPCGLIQH	475	
Db	428	LIGTKKKKFEKLLKSMEEKFTFKVAVRVFNAPLAHQMGAGADLLAVTSRREPCGLIQH	487	
Qy	476	AMHYGTVPVASTGGLVDTVEGVTGFHMGALNPD--KLDEADADALAAATVRRASEVEFAG	533	
Db	498	GMYGTGTCVCASTGGLVDTIVVEGKTGFHMGRLSDVCNVVEPADYKKVATILKRAKVVGVT	547	
Qy	534	GRYPEWANCITSQDLSWSKPAQKWEGLLIEV	564	
Db	548	PAYOBWKNCMLODSWSKGPAPKNWEDVLLEL	578	

```

RESULT 7
T10906
starch synthase (EC 2.4.1.21) - sweet potato
N:Alternate names: starch synthase
C:Species: Ipomoea batatas (sweet potato)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Aug-2002
C:Accession: T10906
R:Wang, S.J.; Yeh, K.W.; Tsai, C.Y.
submitted to the EMBL Data Library, December 1995
A:Reference number: Z17212
A:Accession: T10906
A:Status: preliminary; translated from GH/EMBL/DDDJ
A:Molecule type: mRNA

```

A;Residues: 1-608 <WAS>
A;Cross-references: EMBL:U44126; NID:g1172158; PID:g1172159
A;Experimental source: cv. Tainong; tuberous root
C;Genetics:
A;Gene: SS67
C;Function:
A;Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producing a
A;Pathway: starch biosynthesis
C;Superfamily: starch synthase
C;Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 38.9%; Score 1417; DB 2; Length 608;
Best Local Similarity 51.9%; Pred. No. 1.3e-76;
Matches 299; Conservative 62; Mismatches 171; Indels 44; Gaps 11;

QY 32 LRELARGARKTSRSVAVTGATGATC---ALDVMVAEEVAPWSTGGGLDVTGGPIEL 88
Db 52 LRTSARNLAKMEKMRVQAGTIVCKQOQGMNLFVCGCEGPKCTGGGLDVLGGLPAL 111
QY 89 VKRGHVMIAIPRYDOYADAWDTSVVD--IMG---EKVRYFHSIKKGVHVRWIDHPWFLA 144
Db 112 AARGHVMIVCPRYDQYKDAWETCVVVEPVQGVDRIEPVRFFHSYKRGVDRVFDHPMELE 171
QY 145 KWKGTGSKLYGPRSGADYLDNHRKALFCKAAIEAARVLPF-----GP-CEDCVFVA 196
Db 172 KWKGTGSKLYGPKAGKDYKDNQRLSLLCOAALEAPRVNLNLSKYFSGPYGEDVFFVA 231
QY 197 NWHSLVLPVLKDEYQPKQFTAKSVLAIHIAFQGRMWEBAFDTKLPOAAFDKLA 256
Db 232 NWHHTALLCYLKTMYQSRGIYMNKAVAFCIHNIAYOGRFAFSDFSLNLPDEYKGSFDF 291
QY 257 SDGYAKVYTEANPMEDEKPLTGTYYKINLWLGKGIADKLVTVSPNVAETIAADAAG 316
Db 292 IDGYDK-----PVKG---RKINMKAGIREADRVFTVSPNYAKELVSCVSK 334
QY 317 GVLEDTVIRAKGLEGVINGMDIEWNPKTDKLSAPYDONSVYAGKAAKEALQAEGLP 376
Db 335 GVLEDNHIRDCGIGTCMGMDTQEWNPATDKYLAVKYDITVYMQAPKLKEALQAAVGLP 394
QY 377 VDPAPLAFAGLEBEQKGVDIILALPKILATPKVQVAILGTGKAAEKLNAIGTKYK 436
Db 395 VDRNPILGIFGLBEQKGVDIILALPKILATPKVQVAILGTGKAAEKLNAIGTKYK 453
QY 437 GRAKGVKFSAPLAHMLTAGADFMVPSRFEPCGLIQLHAMHYGTVPVASTGGGLVDTYK 496
Db 454 DKARGVAKENVPLAHMITAGADFMVPSRFEPCGLIQLHAMRYGTPTCASTGGGLVDTYK 513
QY 497 EGVTFHMGALNPD--KLDEADADALATVRRASEVFAGGRYPPEMVANCISQDLSWKP 554
Db 514 EGYTFHMGAFNWDCTVPDVLKVTITVGRALAIYGLAFTEMIKNCMSQELSKGPA 573
QY 555 OKWEGIL-----EEVYVGKGVATAKKEIKVP 592
Db 574 KWEIVLLSLGVAGSEPGVEGE-ETAPLAKENVATP 608

RESULT 8
F86453
granule-bound starch synthase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: F86453
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.B.; Haulzar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: F86453
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-610 <STO>
A;Cross-references: GB:AE005172; NID:g6910568; PID:AAF31273.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: starch synthase

Query Match 38.6%; Score 1405.5; DB 2; Length 610;
Best Local Similarity 49.6%; Pred. No. 6.3e-76;
Matches 299; Conservative 76; Mismatches 163; Indels 65; Gaps 14;

QY 3 VASTSRPSSARPIVINAASFGVKTKANTQALLRELARGARKTSRSVAVTGATGATC--ALD 60
Db 49 VDNLQRSQAKPV-----SAKSKRSKSVKTAGIVCEKGM 85
QY 61 IVMVAEAPWSTGGGLDVTGGLPIELVKRHRVMTIAPRYDOYADAWDTSVVVDI--WG 119
Db 86 VIFGAEGVPWSTGGGLDVLGGLPPALAAARGHRVMTICRYDOYKDAWDTCVVQIKVG 145
QY 120 EK---VRYFHSIKKGVHVRWIDHPWFLAKVWGTGSKLYGPRSGADYLDNHRKALFCKA 176
Db 146 DKVENVRFFHCYKRGVDRVFDHPFIPLAKVWGTGSKLYGPRSGADYLDNHRKALFCKA 205
QY 177 AIEAARVLPF-----GP-CEDCVFVANDHVSALVPLLKDEYQPKQFTAKSVLAIH 228
Db 206 ALEAPQVNLNLSKYFSGPYGEDVFFVANDHHTALLPCYLKSMYQSGVYTNNAKVPCIH 265
QY 229 NIAFGQRMWEEAFKDKLP---QAAFDKLAESDGYAKVYTEATPMEDEKPLTGTYYK 285
Db 266 NIAYGREFAPDDYSLNLPISFKSFD---FMDGYEK-----PVKG---RK 305
QY 286 INWLKGLIADKLVTVSPNVAETIAADAAGVVELTVIRAKGLEGVINGMDIEWNPKT 345
Db 306 INWKAAILEAHVRLTVSPYIAQELISGVDRVELHKLKMTKVTSGGLINGMDYQEWNPST 365
QY 346 DKFLSAPYDONSVYAGKAAKEALQAEGLPDPPTAPLFAFAGLEBEQKGVDIILALPK 405
Db 366 DKYIDIKYDITVTDAKPLKEALQAAVGLPVRDVPVIGFVIGLEBEQKGVDIILALPK 425
QY 406 ILATPKVQVAILGTGKAAEKLNAIGTKYKGRAGKVVKPSAPLAHMLTAGADFMVPSR 465
Db 426 FMGL--NVQMVLITGKTKKMEQAQILEEKPFGKAVGVAKENVPLAHMITAGADFMVPSR 484
QY 466 FEPGGLIQLHAMHYGTVPVASTGGGLVDTYKGVTFHMGALNP--DKLDEADADALAT 523
Db 485 FEPGGLIQLHAMRYGTVPVASTGGGLVDTYKGVTFHMGALNP--DKLDEADADALAT 544
QY 524 VRRASEVFAGGRYPPEMVANCISQDLSWSPKPAQKWEGLLEEVYVGKGVATAKKEIKVP 583
Db 545 VTRAVAVYGTSAQEMVKNMCDQDFSWKGPALWEKVLVSLNVA--GSEAGTEGEET--APL 602
QY 584 AEK 586
Db 603 AKE 605

RESULT 9
T14731
N;Alternate names: starch synthase, granule-bound - sorghum
C;Species: Sorghum bicolor (sorghum)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002
C;Accession: T14731
R;Hsing, Y.C.
submitted to the EMBL Data Library, April 1995
A;Reference number: Z18173
A;Accession: T14731
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-608 <HSI>
A;Cross-references: EMBL:U23945; NID:g1255713; PID:g1255714

A: Experimental source: strain 12311; young seed

C: Genetics:

A: Genome: nuclear

A: Note: wx

C: Function:

A: Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing a

A: Pathway: starch biosynthesis

C: Superfamily: starch synthase

C: Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltransf

P: 1-77/Domain: transit peptide (amyloplast) #status predicted <TNP>

F: 78-608/Product: UDPglucose-starch glucosyltransferase #status predicted <MAT>

Query Match 37.8%; Score 1376; DB 2; Length 608;

Best Local Similarity 46.7%; Pred. No. 3.6e-74;

Matches 295; Conservative 83; Mismatches 180; Indels 74; Gaps 14;

QY 1 MAVASTRPSARPLVINAASFGV-----KKTANQLRLARGARKSTERSAVTGA-- 52

DB 1 MSTLATSQ-----LVATHAGLVGDASMRFRGGVGLRAARAASAAAGDALSMRTSACP 54

QY 53 -----TGATCALTDIVMAAEVAPWSKGTGGLGVDVGTGPIELVKR 91

DB 55 APROQPAARRGGRGFRFSLVVCATAGMNVVFGAEMAPWSKGTGGLGVDVGLGPPAMAAN 114

QY 92 GHRVMTIAPRYDQYADAWDTSVVVDI-MG-----EKVRYFHSIKKKGVHRVWIDHPWFLAKVW 147

DB 115 GHRVNVSPRYDQYKDAWDTSVVSEIKMGDGYETVRFHFCHYKRGVDRVFDHPLFLERVW 174

QY 148 KTGSKLPGSPGADYLDNHKRFALFCKAAATEAARVL-----PF-GP-GEDCVFVANDW 199

DB 175 GKTEEKIYGPAGTDYKQNRFLSLCQAALAPRILSLNNPFGSPGYGDDVVFVNDW 234

QY 200 HSAVLPVLLKDBEQYQKQFTAKSVLAHNTAFQGRMWEAPKDTKLQPAAPDKLAFSDG 259

DB 235 HTGPLSCYLKSNYQSNGLYKDAKTAFCIHNIYQGRFAFSDPPELNLPERFKSSPFDIDG 294

QY 260 YAKVYTEATPMEDEKPLTGTYKKINWLKGGITIAADKLVTVSPNYATEIAADAAGVE 319

DB 295 YEK-----PVGE---RKINNMKAGILEADRVLTVPYAFELISGARGCE 337

QY 320 LDTVIRAKGIGIVNGMDIEENPKTKDFLSAPYDONSYYAGKAALQALQELGVPDP 379

DB 338 LDNIMRLTGITIVNGMDVSEWDPSPKQYIAVKYDVSTAVEAKALKEALQAEVGLPDR 397

QY 380 TAPLFAFGRLEQKQGVDTIILALPKILATPKVQIAILGTGKAAEKLVAIGTYKGRA 439

DB 398 KIPLVAFGRLEQKQGVDMARAIP-LLMEEDIQIVLLGTGKKFERMLMSAEKYPDKV 456

QY 440 KGVVKSAPLAHMLTAGADFMVPSRFPCEGLIQLHAMHYGTVPVVASTGGLVDVYKEGV 499

DB 457 RAVVKFNAALAHIMAGADLLAVTSRFEPCEGLIQLQGMRYGTPCACASTGGLVDVTIEGK 516

QY 500 TGFHMCALNPD--KLDEADALATVRRASEVAGGRYPYPMVANCISODLSWSKPAOKW 557

DB 517 TGFHMRUSVDCNVPEADVKKVATLTKRAIKVGTPTAYEEMVKNCMIQDLSWKGPKNW 576

QY 558 EGLLEE--VYVYKGG---VATAKKEEIKVP 582

DB 577 ENVLLSLGVAGGEPGEGIEGEIAPLAKENVAAP 608

RESULT 10

YUWYU

glycogen(starch) synthase (EC 2.4.1.11) precursor - wheat

N: Alternate names: starch synthase

C: Species: Triticum aestivum (common wheat)

C: Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002

C: Accession: S16261; S33636

R: Clark, J.R.; Robertson, M.; Ainsworth, C.C.

Plant Mol. Biol. 16, 1099-1101, 1991

A: Title: Nucleotide sequence of a wheat (Triticum aestivum L.) cDNA clone encoding the w

A: Reference number: S16261; MUID: 91322506; PMID: 1863765

A: Accession: S16261

A: Molecule type: mRNA

A: Residues: 1-615 <CLA>

A: Cross-references: EMBL:X57233; NID:g21901; FIDN:CAA40509.1; PID:g21902

R: Ainsworth, C.; Clark, J.; Baldson, J.

Plant Mol. Biol. 22, 67-82, 1993

A: Title: Expression, organisation and structure of the genes encoding the waxy protei

A: Reference number: S33636; MUID: 93271462; PMID: 8499619

A: Accession: S33636

A: Molecule type: protein

A: Residues: 71-75, 'X', 77-78 <AIN>

C: Genetics:

A: Gene: waxy

C: Function:

A: Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producin

A: Pathway: starch biosynthesis

C: Superfamily: starch synthase

C: Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltra

P: 1-70/Domain: transit peptide (amyloplast) #status predicted <TNP>

F: 71-615/Product: UDPglucose-starch glucosyltransferase #status experimental <MAT>

Query Match 37.7%; Score 1372; DB 1; Length 615;

Best Local Similarity 49.2%; Pred. No. 6.2e-74;

Matches 279; Conservative 77; Mismatches 159; Indels 52; Gaps 8;

QY 32 LRELARGARKSTSR-----SAVTGATCATCALDIVMAAEVAPWSKGTGGLGVDV 82

DB 42 MFTVGASAPKOSRKPRHFRDRCLSMVVRATCSG-GMNLVFGAEMAPWSKGTGGLGVDV 100

QY 83 GLPIELVKRGHRVMTIAPRYDOYADAWDTSVVVDI-----MGEKRVYFHSIKKVHRVWID 138

DB 101 GLPMAAANGHRVWISPRYDQYKDAWDTSVSEIKVVDYERVRFCHYKRGVDRVFD 160

QY 139 HPWFLAKVWGTGSKLYGPRSGADYLDNHKRFALFCKAAAEARVLPF-----GP---- 188

DB 161 HPCFLEKVRGKTKETIYGPDACTDYEDNQRFSLCQAALFVPRILDLNNPHFSGPYAM 220

QY 189 -----GECVVFANDWHSALVPVLLKDBEQYQKQFTAKSVLAHNTAFQGRMWE 239

DB 221 LCRVPRRAGEDVVFVNDWHTGLLACYLKSNSYQSNGLYRTAKVAFCHINISYQGRFSD 280

QY 240 AFKDTKLQPAAPDKLAFSDGYAKVYTEATPMEDEKPLTGTYKKINWLKGGITIAADKL 299

DB 281 DFAQLNLPORFSSPFDIDGDK-----PVGE---RKINNMKAGILQADKV 323

QY 300 VTVSPNYATEIAADAAGVVELDTVIRAKGIGIVNGMDIEENPKTKDFLSAPYDONS 359

DB 324 LTVSPYAEELISGEARGCELDNIMRLTGITIVNGMDVSEWDPPIKDKELTVNYDVTAL 383

QY 360 AGKAAKAEALQALQELGVPDTPAPLFAFGRLEQKQGVDTIILALPKILATPKVQIAILGT 419

DB 384 EGKALNKEALQAEVGLPVDKVPVLAFAFGRLEQKQGVDMIAAIPVKEEDVQIVLLGT 443

QY 420 GKAAVEKLVNAIGTYKKGRAKGVVKEFSAPLAHMLTAGADFMVPSRFPCEGLIQLHAMHY 479

DB 444 GKXKPELLKSVEEKPTKVRVVRNAPLAHQMGADVLAVTSRFEPCEGLIQLQGMRY 503

QY 480 GTVPVAVASTGGLVDVTVKEGTVGFHMCALNPD--KLDEADALAAVRRASEVAGGRYP 537

DB 504 GTPCACASTGGLVDVTIVEGTGTFHMRUSVDCNVPEADVKKVVTTLTKRAVKVYGTAYH 563

QY 538 EMVANCISODLSWSKPAOKWEGLEEV 564

DB 564 EMVKNCMIQDLSWKGPKNWEDVLEL 590

RESULT 11

S07314

glycogen(starch) synthase (EC 2.4.1.11) precursor - maize

N: Alternate names: starch synthase; UDP-glucose starch glycosyltransferase, starch g

C: Species: Zea mays (maize)

C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 02-Aug-2002

C: Accession: S07314

R: Kloesgen, R.B.; Gierl, A.; Schwarz-Sommer, Z.; Saedler, H.

Mol. Gen. Genet. 203, 237-244, 1986
A:Title: Molecular analysis of the waxy locus of Zea mays.
A:Reference number: S07314
A:Accession: S07314
A:Molecule type: DNA
A:Residues: 1-605 <KLO>
A:Cross-references: EMBL:X03935; NID:g22509; PIDN:CAA27574.1; PID:g1644339
A:Experimental source: line C
A:Note: translation of the nucleotide sequence is not complete
A:Note: part of this sequence, including the amino end of the mature protein, was confirmed by sequencing.
C:Genetics:
A:Gene: waxy
A:Introns: 107/3; 134/3; 167/3; 197/3; 219/1; 252/3; 289/2; 370/3; 431/3; 495/3; 524/3;
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
C:Superfamily: starch synthase
C:Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glycosyltransferase;
F;1-72/Domain: transit peptide (amyloplast) #status predicted <NP>
F;73-605/Product: glycogen (starch) synthase #status experimental <NP>
Query Match 37.6%; Score 1369.5; DB 1; Length 605;
Best Local Similarity 46.7%; Pred. No. 8.6e-74;
Matches 294; Conservative 87; Mismatches 177; Indels 71; Gaps 16;
QY 1 MAVASTSPSSARPPIVINAASFGVKTANOLLRELARG--SARKSTSRSAVTCATGA--- 55
Db 1 MAALATSQ-----LVATRAGLV-PDASTFRGAAGGLRGARASAAADTLMSRTSARAA 53
QY 56 -----TCA---LDIVMAAEVAPWSKTGGGLDVTGGPLIELVKRGHRV 95
Db 54 PRHQOARRGRFPVSLVVCASAGMNVFVGAEMAPWSKTGGGLDVLGGLPPMAAANGHRV 113
QY 96 MTIAPRYDQADAWTSSVVDI-MG---EKVYFHSIKGVHVRVTHDPWFLAKVWGKTG 151
Db 114 WYVSPRYDQKAWDTSSVSEIKMGDGYETVFFHCYKRGVDRVFVDHFLERVWGKTE 173
QY 152 SKLYGPRGADYLDNHKRFALPKAAIEAARVL-----PF--GP-CEDCVFEVANDWHSAL 203
Db 174 EKIIYGVAGTDYRDNQLRSLCQALAPRILSLNNNPFSPGYEDVFFVNCNDWHTGP 233
QY 204 VVILLKDEYQPGQRTKAKSVIAIHNIAPQGRMWEAEFKDTKLPOAAFDKLAFLSDGYAKV 263
Db 234 LSCYLKSNYSQSHGIYRDATKATPCIHNIYQGRFAFSYDPELNLPERKSSDFIDGYEK- 292
QY 264 YTEATPMEDEXPPDTGTYKIKNLKGLIADKLIVTSPNYATEIADADAGGVLDIV 323
Db 293 -----PVEG---RKINWKGILLEADRVLTIVSPYAEELISGIAIGCELDNI 336
QY 324 IRAKGIETVINGMDIEENPNKTDKFLSAPYDQNSVYAGKAAKALQALQELGLPVDPTAPL 383
Db 337 MRLTGITVINGMDYSEWDPSRDKYIAVYDYSTAVEAKALNKEALQAEVGLPVDNRNPL 396
QY 384 FAFIGRLSEQKVDIILALPKIL-ATPKVQTAIIGTGKAAVEKLVNATIGTKYGRAGV 442
Db 397 VAFIGRLSEQKVDVYMAAIPOLMEMVEDVQIVLLGTGKKEPERMLMSAEKFPCKVRV 456
QY 443 VKFSAPLAHMLTAGADMLVPSREPCGLIQHMHYCTVPVAVSTGGGLVDIVKESVTF 502
Db 457 VKENALAHMHMAGADVLAVTSRFEPCGLIQGMRYGTGTPCACSTGGGLVDIVIEGRTGF 516
QY 503 HMGALNPD--KLDEADADALATVRRASEVAGGRYPPEMVANCISODLSWSKPAQKWEGL 560
Db 517 HMGRLSVCNVVEPADVKVATTLQRAIKVVGTPAYEEMVRCMIDQLSWKGPARNWENY 576
QY 561 LEE--VYVKGKV-----ATAKKEIKVP 582
Db 577 LLSLGVAGGEPGVEGEIAPLAKENVAAP 605
RESULT 12
T07921
probable starch synthase (EC 2.4.1.21), granule bound - Chlamydomonas reinhardtii (fragm
N:Alternate names: probable starch synthase

C:Species: Chlamydomonas reinhardtii
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002
C:Accession: T07921
R:D'hulst, C.; Abel, G.; Kossmann, J.; Ball, S.
submitted to the EMBL Data Library, September 1997
A:Description: Cloning of cDNAs coding for starch synthases in the green algae Chlam.
C:Reference number: Z16218
A:Accession: T07921
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-238 <DHU>
A:Cross-references: EMBL:AF026420; NID:g3169784; PID:g3169785
C:Genetics:
A:Gene: STA2
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producti
A:Pathway: starch biosynthesis
C:Superfamily: starch synthase
C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase
Query Match 32.8%; Score 1194; DB 2; Length 238;
Best Local Similarity 99.2%; Pred. No. 6.3e-64;
Matches 234; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 473 QLRHMYGTVPVVASTGGLVDIVKEGVTGFHMGALNPKLDEADADALATVRRASEVFA 532
Db 3 RVHMYGTVPVVASTGGLVDIVKEGVTGFHMGALNPKLDEADADALATVRRASEVFA 62
QY 533 GGRYPPEMVANCISODLSWSKPAQKWEGLLEEVYVYGGVATKKEIKVPVAKIPGDLP 592
Db 63 GGRYPPEMVANCISODLSWSKPAQKWEGLLEEVYVYGGVATKKEIKVPVAKIPGDLP 122
QY 593 AVSYAPNTLKPVSASVEGNGAAAPKVGTTAPAMGAWRATTPSGPSPAAATPKVTTKPAL 652
Db 123 AVSYAPNTLKPVSASVEGNGAAAPKVGTTAPAMGAWRATTPSGPSPAAATPKVTTKPAL 182
QY 653 PATAKPKTAGLKLAGEASTTSTSENGAASNGNGASAKTSAAKPLVSAATKSA 708
Db 183 PATAKPKTAGLKLAGEASTTSTSENGAASNGNGASAKTSAAKPLVSAATKSA 238
RESULT 13
T07668
starch synthase (EC 2.4.1.21) SSI precursor - potato
C:Species: Solanum tuberosum (potato)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002
C:Accession: T07668
R:Abel, G.J.W.
submitted to the EMBL Data Library, January 1997
A:Reference number: Z16081
A:Accession: T07668
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-641 <ABE>
A:Cross-references: EMBL:Y10416
A:Experimental source: cv. Desiree; leaf
C:Genetics:
A:Gene: SSI
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose productin
A:Pathway: starch synthase
C:Superfamily: starch synthase
C:Keywords: glycosyltransferase; hexosyltransferase
Query Match 24.8%; Score 903; DB 2; Length 641;
Best Local Similarity 38.1%; Pred. No. 4.8e-46;
Matches 219; Conservative 91; Mismatches 183; Indels 82; Gaps 20;
QY 30 QLLRELARGSARKSTSRSAVTG-----ATGATCALDIVMAAEVAPWSKTGG 76
Db 89 QLIPHSVAGDATWVESHDIIVANDRDLSEDTEEMETPIKLTFFNIIFVTAAAPSKTGG 148
QY 77 LGDVTGGPLIELVKRGHRVMTIAPRY-----DOYADAWDTSV--VVDIMG--EKVRYF 125

[illegible]

RESULT 14
S61505
glycogen(starch) synthase (EC 2.4.1.11) isoform II precursor - garden pea
N:Alternate names: glycogen(starch) synthase (EC 2.4.1.11) isoform II precursor; granule
C:Species: Pisum sativum (garden pea)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 16-Aug-2002
C:Accession: S61505; S72373; S72312
R:Dry, I.; Smith, A.; Edwards, A.; Bhattacharyya, M.; Dunn, P.; Martin, C.
Plant J. 2, 193-202, 1992
A:Title: Characterization of cDNAs encoding two isoforms of granule-bound starch synthase
A:Reference number: S61504; MUID:93251108; PMID:1302049
A:Accession: S61505
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-752 <DRY>
A:Cross-references: EMBL:X88790
A:Accession: S72373
A:Molecule type: protein
A:Residues: 58-59, 'H', 61-73 <DRW>
K:Edwards, E.A.
submitted to the EMBL Data Library, June 1995
A:Reference number: S72312
A:Accession: S72312
A:Molecule type: mRNA
A:Residues: 1-85, 'KVLAQRELIQIQIAEKK', 104-139, 'SSSSGSNAVETKRWHCQQQLC', 160-752 <EDW>
A:Cross-references: EMBL:X88790; NID:9887573; PID:9887573
C:Keywords: glycosyltransferase; hexosyltransferase
F:1-57/Domain: signal sequence #status predicted <SIG>
F:59-752/Product: glycogen (starch) synthase isoform II #status experimental <MAT>

63	QY	MVAEEVAPWSKTGGLGVDVTGGLPLELIVKRGHFVMTAPRYDQYADAWDTSV--VVDIMGE	120
265	Db	LVSEACAPWSKTGGLGVDVAGSLPKALARRGHRFVMI VAPHYGTAEAHDTGVKRYKVAGQ	324
121	QY	--KYRYFHSIKKGVRHVIWDHPWFLAKVWCKTSGSKLYGPRSGADYLDNKHRELFALCKAAI	178
325	Db	DMEVTFYTHYIDGVDFVFIOSP-----IFRNLESNIY-----CGNRDLDIRRMVLFCKAAV	375
179	QY	EAARVLP-----FGPGEDCVFVANDWHSALVPVLLKDEYOPKGOFTKAKSVLAIRNIAFP	233
376	Db	EVPHVPCGGGICYGDG--NLVFIANDWETALLPVYLKAYYRDHGLMAYTRSVLVIHIIAHQ	434
234	QY	GRMDEEAFKDKLPQAAFPDKLATSDGVAKYVYTTATPMEEDEKPPLTGKTYKKINLWLGKI	293
435	Db	GRG-----PVEDENTVDLSNGYLDLFLKMYDPVGGEH-----FNIFPAAGL	473
294	QY	IAADKLTVTSPNYATEIAADAAGGVGLDVTYIRAKG--LEGIVNGMDIEEWNPKTKFLSA	351
474	Db	KTADRIVTVSUGYAWELKT--SEGQWGLHIIINESDWKFRGIVNGVDTKDNPFQDAYLTS	532
352	QY	P-----YDQNSYYAGKAAAKAALQALQELGDPVDTAPLPAFTGRLEEOQGVDDIILALPKIL	407
533	Db	DGYTNYNLKTLQTKRQCKAALQRELGLPVEDVPFIISFTGRLDHQGVGLDIAEAIPWMW	592
408	QY	ATPKVQIAILGTGKAAYEKLVAIGTKYKGRAGGVVKFSAPLAHMLTAGADEPMLVPSREE	467
593	Db	--SHDVQIVMLGTGRADLEQMLKEPFAQCHDK--RSWVGFSVKMAHRIITAGSDILLMPSPRE	651
468	QY	PGGLTOLHAMHYGVTPVYVASTGGVLVDIVK-----EGVTGFHMGALNPKDLDEADADALA	521
652	Db	PGGLNQLVAMSYGVTPVVHVGUGLUDIVQGFENPFDESGVGW-----TFDRAEANKLM	703
522	QY	ATVRRASEVFAGGRYPENVANC-----SODLSWSKPAQKWEGLLEE	563
704	Db	AA-----LWNCLLITYDKYKSNBEGIOERGMSQDLSWDNAAQOQVEVLVA	747
564	QY	VVY 566	
748	Db	AKY 750	
RESULT 15			
QJ2322		starch synthase (EC 2.4.1.21) precursor - rice	
		N:Alternate names: starch synthase	
		C:Species: Oryza sativa (rice)	
		C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 16-Aug-2002	
		C:Accession: JQ2322; FQ0811	
		R:Rabe, T.; Nishihara, M.; Mizuno, K.; Kawasaki, T.; Shinada, H.; Kobayashi, E.	
		Plant Physiol. 103, 565-573, 1993	
		A:Title: Identification, cDNA cloning, and gene expression of soluble starch	
		A:Reference number: JQ2322; MUID:94302151; PMID:7518089	

A;Residues: 1-626 <BA61>
A;Cross-references: DDBJ: D16202; NID: g450484; PID: g450485
A;Accession: P00811
A;Molecule type: protein
A;Residues: 114-129 <BA62>
A;Experimental source: seed
C;Function:
A;Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose
C;Superfamily: starch synthase
C;Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltran
F;114-626/Domain: transit peptide (amyloplast) #status predicted <TRS>
F;114-626/Product: ADPglucose-starch glycosyltransferase, 57KD form #statu
F;122-626/Product: ADPglucose-starch glycosyltransferase, 55kd form #statu
F;440-448/Region: substrate binding #status predicted

Query Match 24.0%; Score 873.5; DB 2; Length 626;
Best Local Similarity 41.5%; Pred. NO. 2.6e-44;
Matches 197; Conservative 77; Mismatches 154; Indels 47

QY 61 IYVVAEYAPWSKTGGLGVDTGGLPIELVYKGRHVRMTIAPRY-----DOYADAWTSTSV 114
Db 135 VVFTGEASPYAKSGGLGVCGSLPTALALRGHRVNVVMRYMNGALNKNFANAFYTEKH 194
QY 115 VDIM---GE-KVRYFHSIKKGVRHWIDHPWFLAKVWGKTGSKLYGPRSCADYLDNHRF 170
Db 195 IKIPCFGGEHEVTFEYHEYRDSVDWVFDHPSY-----HRPGNLYGDNFGA-FGDQOFY 247
QY 171 ALFCRAAIEAARVLPFGP---GEDCVFVANDWHSALVPVLLKDEYQPKQFTKAKSVLAI 227
Db 248 TLLCYAACAPLILELGGYIYGCKMVFVNDWHSALVPVLLAARYPYGVYRDARSVIVI 307
QY 228 HNTAFQGRMWEFAFKDTKLPOAFDKLAFSDGYAKYITEATPMEDEKPPPLGKTYKKIN 287
Db 308 HNLHQGVPEASTYPLGLPPEWYGALEW-----VFPEWARRHALDKG-----EAVN 354
QY 288 WLKGGITAAADKLVTSPNYATEIAADAAGGVDELDTVI--RAGIEGIVNGMDIEWNPKT 345
Db 355 FLKGAVVTADRIVTYSQGYSEWVTT-AEGGQGLNELLSSRKSVLNGIVNGIDINDWNPT 413
QY 346 DKFLSAPYQNSVYAGKAAKALQALGLPVDPTAPLFAFTGRLEEOKGVDIILAALEPK 405
Db 414 DKELPYHYSVDDL-SGKAKCAELQELGLPIRPDVPPLIGFTGRLDYQKGLDLKLAIPD 472
QY 406 ILATPKVQJAILGTGKAAVEKLVNAIGTKYKGRKGVVKSAPLAHMLTAGADFMVLPSP 465
Db 473 LM-RDNIQVFMVLSGDPGEGWMRSTESGYRDKFRGWGFSVPVSHRITAGCDILLMPSP 531
QY 466 FEPCGLIQLHAMHYGTVPVVASTGGGLVDTVK-----EGVTGFHMGALNPKD 511
Db 532 FEPCGLNQLYAMQYGTVPVVGHTGGLRDIVENFNPFAEKGEQGTGWAFSPLTIEK 586

Search completed: June 4, 2003, 14:57:22
Job time : 24.3001 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:50:41 ; Search time 9.50655 Seconds
(without alignments)
3088.950 Million cell updates/sec

Title: US-09-980-771a-3
Perfect score: 3641
Sequence: 1 MAVASTSRPSSARPIVINAA.....SASKISAAKPLVSAATRKSA 708

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1485.5	40.8	608	1	UGST_MANES
2	1466	40.3	608	1	UGST_ANTMA
3	1463.5	40.2	607	1	UGST_SOLTU
4	1431.5	39.3	603	1	UGST_PEA
5	1426	39.2	609	1	UGST_ORYSA
6	1423	39.1	609	1	UGST_ORYGL
7	1421.5	39.0	603	1	UGST_HORVU
8	1417	38.9	608	1	UGST_IPOBA
9	1376	37.8	608	1	UGST_SORBI
10	1372	37.7	615	1	UGST_WHEAT
11	1369.5	37.6	605	1	UGST_WAIZE
12	903	24.8	641	1	UGS2_SOLTU
13	874	24.0	752	1	UGS2_PEA
14	873.5	24.0	626	1	UGS2_ORYSA
15	869	23.9	610	1	UGS2_WHEAT
16	849.5	23.3	788	1	UGS3_SOLTU
17	684.5	18.8	477	1	GLGA_STRPN
18	677.5	18.6	484	1	GLGA_BACST
19	670.5	18.4	485	1	GLGA_LACLA
20	669	18.4	478	1	GLGA_LACLA
21	653.5	17.9	480	1	GLG1_RHIME
22	652	17.9	480	1	GLGA_AGR5
23	651	17.9	486	1	GLGA_THEMA
24	646.5	17.8	486	1	GLGA_RHTR
25	637	17.5	477	1	GLGA_CLOAB
26	622	17.1	482	1	GLGA_CLOPE
27	622	17.1	484	1	GLGA_VIBCH
28	614.5	16.9	476	1	GLGA_BACHD
29	604.5	16.6	486	1	GLG2_RHIME
30	604	16.6	476	1	GLGA_YERPE
31	602.5	16.5	481	1	GLGA_RHILO
32	587	16.1	477	1	GLGA_ECOLI
33	587	16.1	477	1	GLGA_SALTI

34	580	15.9	477	1	GLGA_SALTY
35	578	15.9	480	1	GLGA_PASMU
36	570.5	15.7	1230	1	UGS4_SOLTU
37	569	15.6	476	1	GLGA_HAEIN
38	560.5	15.4	472	1	GLGA_ANASP
39	555	15.2	461	1	GLGA_FUSNN
40	552.5	15.2	463	1	GLGA_AQUAE
41	538	14.8	465	1	GLGA_SYN7
42	536	14.7	477	1	GLGA_RHOSH
43	536	14.7	492	1	GLG2_ANASP
44	535	14.7	477	1	GLGA_SYNY3
45	527	14.5	444	1	GLGA_DEIRA

ALIGNMENTS

RESULT 1
UGST_MANES STANDARD; PRT; 608 AA.
AC Q43784;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Granule-bound glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11).
GN WAXY OR GBSS.
OS Manihot esculenta (Cassava) (Manioc).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Malpighiales; Euphorbiaceae; Manihot.
OX NCBI_TaxID:3983;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. M.COL.22; TISSUE=Tuberous root;
RX MEDLINE=94083565; PubMed=8260633;
RA Salehuzzaman S.N., Jacobsen E., Visser R.G.F.;
RT "Isolation and characterization of a cDNA encoding granule-bound
RT starch synthase in cassava (Manihot esculenta Crantz) and its
RT antisense expression in potato.";
RL Plant Mol. Biol. 23:947-962(1993).
CC -!- FUNCTION: RESPONSIBLE FOR THE SYNTHESIS OF AMYLOSE IN RESERVE
CC STARCH.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =
CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).
CC -!- PATHWAY: Starch biosynthesis.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN A NUMBER OF DIFFERENT ORGANS,
CC BUT MOST ABUNDANTLY IN TUBERS.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: X74160; CAA52273.1;
InterPro: IPR001296; Glycos_transf_1.
Pfam: PF00534; Glycos_transf_1; 1.
GlycoGen biosynthesis; Transferase; Glycosyltransferase;
KW Transit peptide; Chloroplast; Starch biosynthesis.
FT CHAIN 79 608 CHLOROPLAST (BY SIMILARITY).
FT BINDING 96 96 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
SQ SEQUENCE 608 AA; 66968 MW; C9C970CD3011BDBD CRC64;

Query Match 40.8%; Score 1485.5; DB 1; Length 608;
Best Local Similarity 53.6%; Pred. No. 1e-80;
Matches 304; Conservative 67; Mismatches 157; Indels 39; Gaps 11;


```
Query Match 39.2%; Score 1426; DB 1; Length 609;
Best Local Similarity 48.9%; Pred. No. 3.5e-77;
Matches 302; Conservative 81; Mismatches 172; Indels 62; Gaps 13;

QY 3 VASTSRPSS-----ARPIVINAASFVKKKTAN---QLLRELARGSARKSTSRSA 48
DB 18 IADRSAPSLRLHGFQGLKPRSPAGGDATLSVTTTSARATPKQORSVQGRSRR---FPVS 74

QY 49 VTGATGATCADIWVAEVAEPWSTGGGLDVTGGLPIELVKGHRVMTIAPRYDQYADA 108
DB 75 VVIATGA--GMNVFVGAEMAPWSTGGGLDVLGGLPPAMAANGHRVWVISPRIYDQYKDA 132

QY 109 WDTSVVVDI-----MGEKVYFHSIKGVHRVWIDHPWFLAKVWGKSGKLYGPRSGADYL 164
DB 133 WDTSVVAEIKVADRYERVRFFHCYKRGVDRVIDHPSEFLEKVGWGTGKIYGPDTGVYDK 192

QY 165 DNHRKRALFCKAAIEAARVL-----PFGP---GEDCVFVANDWHSALVPVLLKDEYQPKG 216
DB 193 DNQMFSLLCQAALAPILNINNPNYFKGTGDEYVFCVNDWHGGLPLASYLKNNYQNG 252

QY 217 QFTKAKSVLAIHNIAPQGRMWEAFKDTKLPAQAFDKLAFSDGYAKVYTEATPMEDEK 276
DB 253 IYRNAKVAFCIHNISYQGRFAFEDYPELNLSEFSSPFDIDY-----DTPVEG---- 302

QY 277 PLTGKTYKINWLKGGITAAADKLIVTVSPNYATEIAADAAGGVDELDTVIRAKGIEGIVNGM 336
DB 303 -----RKINWMKAGILEADRLVTVSPYIAELISGIARGCELNIMLTGITIVNGM 355

QY 337 DIEENPKTKFLSAPYDQNSYAGKAAKALQAEGLPVDPTAPLFAFIRGLREEQKV 396
DB 356 DVSEWDPSKDYITAKYDATTATIAEAKALNKEALQAEGLPVDKRIPLTAFIRGLREEQKV 415

QY 397 DIILAALPKILATPKVQITAILGTGKAAYEKLVNAIGTKYGRKAKGVKFSAPLAHMLTAG 456
DB 416 DVMAAAIPELM--QEDVQIVLLGTGKKFELKLSMEEKYPGKVRVAVVFNAPLAHLMAG 474

QY 457 ADEMLVPSRFPCCGLIQLHAMHYGTVPVVASTGGGLVDTVKGVTGFHMGALNPD--KLDE 514
DB 475 ADVLAIVPSRFPCCGLIQLQGRYGTIPCACTGGGLVDTVIEGTFHMGRLSVDCKVVEP 534

QY 515 ADADALAATVRRASEVFAAGGRYPENVANCISODLSWSKPAQKWEGLLEEVYVYKGG---- 570
DB 535 SDVKKVAATLRAIKVVGTPAYEEMVRNOMDLSWKGPAKNWENVL--LGLGVAGSAPG 592

QY 571 -----VATAKKEIKVP 582
DB 593 IEGDEIAPLAKENVAAP 609

RESULT 6
UGST_ORYGL STANDARD; PRT; 609 AA.
AC Q42968;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Granule-bound glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11).
GN WAXY.
OS Oryza glaberrima (African rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4538;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. GMS1;
RX MEDLINE=92134825; PubMed=1685658;
RA Umeda M., Ohtsubo H., Ohtsubo E.;
RT "Diversification of the rice Waxy gene by insertion of mobile DNA
RL elements into introns.";
RL Jpn. J. Genet. 66:569-586(1991).
CC -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.
```

Query Match	39.0%;	Score 1421.5;	DB 1;	Length 603;
Best Local Similarity	51.0%;	Pred. No. 6.3e-77;		
Matches 291;	Conservative 76;	Mismatches 167;	Indels 37;	Gaps 10;
QY	8	RPSARPVINAASFVKYKTYANQLLRELARGSARKSTSRVATGATGATCALTIVMVAE	67	
		: : : : : : : : : : : : : : :		
Db	31	RPRNPADAALGMRTICASAAPKQ-SRKAHRGSR--CLSVVVSATGS--GMNLVFGAE	84	
QY	68	VAPWSKTGLDVTGGLPIELVKRGHRVMTIAPRYDQYADAWDTSVVVDIM----	GEKVR	123
		: : : : : : : : : :		
Db	85	MAPWSKTGLDYLGLPMAANGHRVWVSPRYDQYKDWDTSVISEIKVADEYER	144	
QY	124	YHSIKKGVHRVWIDHPWFLAKVWGKTCSLKLYGPRSGADYLDNHKRFALFCKAAIEAARV	183	
		: : : : : : : : : : : : : : : : : : : : : : : : :		

FT CHAIN 77 608 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
 FT BINDING 96 96 UDP-GLUCOSE (BY SIMILARITY).
 SQ SEQUENCE 608 AA; 67000 MW; CCA8FA50A2F69CB0 CRC64;

Query Match 38.9%; Score 1417; DB 1; Length 608;
 Best Local Similarity 51.9%; Pred. No. 1.2e-76;
 Matches 299; Conservative 62; Mismatches 171; Indels 44; Gaps 11;

QY 32 LRELARGARKSTSRSAVGTGATC---ALDIVMVAEAVAPWSKTGGDVTGGLPIEL 88
 DB 52 LRTSARNLAKMKRVEQWQAGTIVCKQGMNLFVFGCEGPKCTGGGLGVLGPPAL 111
 QY 89 VKRGHRVMTIAPRYDOYADAWDTSVVD-IMG---EKVRYFHSIKGVHRVWIDHPWFLA 144
 DB 112 AARGHRVMTIAPRYDOYADAWDTSVVD-IMG---EKVRYFHSIKGVHRVWIDHPWFLA 171
 QY 145 KVMGTGSKLYGPRGADYLDNHRKFALECKAAIEAARVLPF-----GP-GEDEVFA 196
 DB 172 KVMGTGSKLYGPRGADYLDNHRKFALECKAAIEAARVLPF-----GP-GEDEVFA 231
 QY 197 NDHSAVPLVLLKDYQPKQGTAKSVLAHINIAFOGRMWEAEAFKDTKLPQAAFKLAF 256
 DB 232 NDHSAVPLVLLKDYQPKQGTAKSVLAHINIAFOGRMWEAEAFKDTKLPQAAFKLAF 291
 QY 257 SDGYAKVYTEATPMEDEKPPLTGTYKINMLKGGIITAADKLVTVSPNVATEIAADAAG 316
 DB 292 IDGYDK-----PVKG---RKINMKAGIREADRVTVSPNAYKELVSCVSK 334
 QY 317 GVLEDTVIRAKGIEGIVNGMDEENPKTKFSLAPYDONSVYAGKAAKALQAEGLP 376
 DB 335 GVLEDTVIRAKGIEGIVNGMDEENPKTKFSLAPYDONSVYAGKAAKALQAEGLP 394
 QY 377 VDPAPLAFAGRLBEQKGVDIILALPKILATPKVQIATGTAAYEKLNAIGTKYK 436
 DB 395 VDRNPLPLGIFGRLEEQKSDILYAAISKFSW-DVQILITGKKKFEQEQLEVMVP 453
 QY 437 GRAKGVKFSAPLAHMLTAGADFMVPSRFPCEGLIQLHAMHYGTVPVVASTGGGLVTVK 496
 DB 454 DKARGVAFNVDLAMIYTAGADFMVPSRFPCEGLIQLHAMHYGTVPVVASTGGGLVTVK 513
 QY 497 EGVTFHMGALNPD--KLDEADADALAAATVRRASEVFAAGRYPEMVANCIQDLSWSKPA 554
 DB 514 EGYTFHMGALNPD--KLDEADADALAAATVRRASEVFAAGRYPEMVANCIQDLSWSKPA 573
 QY 555 QKWEGLL-----EEVYKGGVATAKKEIKVP 582
 DB 574 KMWETVLLSLGVAGSEPGVEGE-EIAPLAKENVAATP 608

RESULT 9

UGST_SORBI
 ID UGST_SORBI STANDARD; PRT; 608 AA.
 AC Q43134;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Granule-bound glycogen [starch] synthase, chloroplast precursor
 DE (EC 2.4.1.11).
 GN WAXY OR WX.
 OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoidae; Andropogoneae; Sorghum.
 OX NCBI_TaxID=4538;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. 12311; TISSUE=Seed;
 RA Hsing Y.C., Liu C., Yu H., Hsieh J.;
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-alpha-D-glucosyl](N) =
 CC UDP + [(1,4)-alpha-D-glucosyl](N+1).
 CC -!- PATHWAY: Starch biosynthesis.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U23945; AAC49804.1; -;
 DR InterPro; IPR001296; Glycos_transf_1.
 DR Pfam; PF00534; Glycos_transf_1; 1.
 KW Glycogen biosynthesis; transferase; Glycosyltransferase;
 FT Transit peptide; Chloroplast; Starch biosynthesis.
 FT TRANSIT 1 77 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 78 608 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
 FT BINDING 97 97 UDP-GLUCOSE (BY SIMILARITY).
 SQ SEQUENCE 608 AA; 66074 MW; C31333FA87D2D8A6 CRC64;

Query Match 37.8%; Score 1376; DB 1; Length 608;
 Best Local Similarity 46.7%; Pred. No. 3.1e-74;
 Matches 295; Conservative 83; Mismatches 180; Indels 74; Gaps 14;

QY 1 MAVASTSRPSSARPPIVINAASFGV-----KKTANQLLRELARGSKSTSRSAVGTGA-- 52
 DB 1 MSTLATSQ-----LVATHAGLVDPASMFRRGGVQGLRAAARASAAAGDALSMRTSACP 54
 QY 53 -----TCATCALDIVMVAEAVAPWSKTGGDVTGGLPIELVKR 91
 DB 55 AFRQQAARRGGRRGRRPFLVVCATAGMNVFVGAEMAPWSKTGGDVTGGLPAPMAAN 114
 QY 92 GHRVMTIAPRYDOYADAWDTSVVVDI-MG---EKVRYFHSIKGVHRVWIDHPWFLAKVW 147
 DB 115 GHRVMTIAPRYDOYADAWDTSVVSEIKMGDGYETVRFHCYKRGVDRVIDHPLFLERVW 174
 QY 148 GKTGSKLYGPRGADYLDNHRKFALECKAAIEAARVLPF-----PF--GP-GEDEVFA 199
 DB 175 GKTGSKLYGPRGADYLDNHRKFALECKAAIEAARVLPF-----PF--GP-GEDEVFA 234
 QY 200 HSALPVLVLLKDYQPKQGTAKSVLAHINIAFOGRMWEAEAFKDTKLPQAAFKLAFSDG 259
 DB 235 HTGPLSCYLKSNFQSGIYKDKATKAFCHINISYQGRFAPSDFPDELNPERFKSSPFDIDG 294
 QY 260 YAKVYTEATPMEDEKPPLTGTYKINMLKGGIITAADKLVTVSPNVATEIAADAAGVE 319
 DB 295 YEK-----PVGE---RKINMKAGILEADRVTVSPNAYEELISGIARGCE 337
 QY 320 LDTVIRAKGIEGIVNGMDEENPKTKFSLAPYDONSVYAGKAAKALQAEGLPVPD 379
 DB 338 LDNIMRLTGITGIVNGMDVSEWDPKDKYIAVDVSTAVEAKALKEALQAEVGLPVD 397
 QY 380 TAPLFAFGRLEEQKGVDIILALPKILATPKVQIATGTAAYEKLNAIGTKYKGR 439
 DB 398 KIPLFAFGRLEEQKGVDMVAAAIP-LLMEEDIQIVLLGTGKKKFERMLMSAEKYPDKV 456
 QY 440 KGVKFSAPLAHMLTAGADFMVPSRFPCEGLIQLHAMHYGTVPVVASTGGGLVTVK 499
 DB 457 RAVVKNALAAHIMAGADLLAVTSFEPCEGLIQLQGMRYGTPCASCATGGGLVTVIEGK 516
 QY 500 TGFHMGALNPD--KLDEADADALAAATVRRASEVFAAGRYPEMVANCIQDLSWSKPAQKW 557
 DB 517 TGFHMGALNPD--KLDEADADALAAATVRRASEVFAAGRYPEMVANCIQDLSWSKPAQKW 576
 QY 558 EGLLEE--VYKGG-----VATAKKEIKVP 582
 DB 577 ENVLLSLGVAGSEPGVEGEIAPLAKENVAAP 608

RESULT 10

UGST_WHEAT

ID UGST_WHEAT STANDARD; PRT; 615 AA.

[illegible]


```
SQ SEQUENCE 605 AA; 65966 MW; 137F15207DBFC189 CRC64;
Query Match 37.6%; Score 1369.5; DB 1; Length 605;
Best Local Similarity 46.7%; Pred. No. 7.5e-74;
Matches 294; Conservative 87; Mismatches 177; Indels 71; Gaps 16;

QY 1 MAVASTRPSSARPIVINAASFGVKKTAQNLRLRLARG--SARKSTSRSAVTGAIGA--- 55
DB 1 MAALATSQ-----LVATRALGV-PDASTFRGAAGLGRASAAAADTILSMRTSRAAA 53
QY 56 -----TCA---LDIVMAAEVAPWSKTGLDVTGGLPGLLELVKRGHRV 95
DB 54 PRHQOQARRGRRFSLVVCASAGNMFVVGAEAPWSKTGLDVTGGLPGLPAMANGHRV 113
QY 96 MTIAPRDQADAWDTSVVVDI-MG---EKVRYFHSIKKGVHVRWIDHPWFLAKWGTG 151
DB 114 MVVSPRDQYKDADMTSVVSEIKMGDGYETVRFFCYKRGVDRVFDVHPLERVWGKTE 173
QY 152 SKLYGPRSGADYLDNHRKFALECKAAFEAAARVL-----PF--GP-GEDCVFVANDHSA 203
DB 174 EKIVGPVAGTYDRNQRLRFLSCQALEAPRLILSLNNNPFSGPYGEDVWFVCDNHTGP 233
QY 204 VPVLKDEYOPKQFTKAKSVLAIHNAFQGRMWEAEFKDTPKLPQAAFDKLAIFSDGYAKV 263
DB 234 LSCYLKSNYQSHGIYRDAKTAFCIHNISYQGRFAFSYPELNLPERKSSDFDIDYEK- 292
QY 264 YTEATPDEEKPLTKTKYKINWLKGGIITAADKLVTVSPNYATEITAADAAGVELDTV 323
DB 293 -----PVGE---RKINMKAGILEADRVLTVPYPAEELISGIARGCELDNI 336
QY 324 IRAKIGIEVINGMDIEWNPKTKFLSAPYDONSVYAGKAAKALQALQELGLPDPAPL 383
DB 337 MRLTGITGVINGMDVSEWDSRDKYIAVKDVTAVAKALNKEALQALQAEVGLPVDNRIPL 396
QY 384 FAFTGRLEQKGVDIILAALPKIL-ATPKVQIAILGTGKAAEYKLVNAIGTKYKGRAGV 442
DB 397 VAFITGRLEQKGVDMVMAAIPOLMEMVEDQVILVLTGKFKKFERMMSAEKFFGKVRV 456
QY 443 VKFSAPLAHMLTAGADPMVPSRPEPCGLIQLHAMHYGVPPVVASTGGLVTVKGVTF 502
DB 457 VKTNAALAHMAGADVLAVTSRPEPCGLIQLGMRYGTPCACASTGGLVTVILEGKTGF 516
QY 503 HMGALNPD--KLDEADADALAATVRRASEVPAGSRYPMPVANCISQDLSNKSQAQKEGL 560
DB 517 HMGELSDVCNVPEADVKVATILQRAIKVGVTPAYEEMVRNCMIQDLSMKGPAPKNWEN 576
QY 561 LEE--VYVGKGV-----ATAKKEEIKVP 582
DB 577 LLSLGVAGGEPGVBGEETIAPLAKENVAAP 605

RESULT 12
UGS2_SOLTU STANDARD; PRT; 641 AA.
AC P93568;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Soluble glycogen (starch) synthase, chloroplast precursor
DE (EC 2.4.1.1) (SS I).
OS Solanum tuberosum (potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Desiree; TISSUE=Leaf;
RX MEDLINE=97164391; PubMed=9011082;
RA Abel G.J.W., Springer F., Willmitzer L., Kossmann J.;
RT Cloning and functional analysis of a cDNA encoding a novel 139 kDa
RT starch synthase from potato (Solanum tuberosum L.);
RL Plant J. 10:981-991(1996).
```

QY	294	IAADKLVTVPENYATETAAADAGVVELDVTVIRAKG--IEGIVNGMDIEENWPKTDKFLSA	351
Db	474	KTADRIVTVSHGYAWELKT-SEGQWGHLHIIINESDKWFRGIVNGVDTKWNPNQDFAYLTS	532
QY	352	P-----YDONSYPYAGKAAKALQAEGLGLPVDPTAPLFAFTGRLEBEQGVDIILAAALPKIL	407
Db	533	DGYTNYNKLITQCKROCKAALQRELGLPVREDYPIISFIRGLDHQKGVDLIAEAIWMM	592
QY	408	ATPKVQIAILGTGKAAYEKLVAITGYKGRAKGVKFSAPLAHMLTAGADFMVLVPSRFE	467
Db	593	-SHDVQLVMLGTGRADLEQMLKEFEAQHCDCIRSVWGVGSFKNAHRIITAGSDILLMPSRFE	651
QY	468	PCGLIQLHAMHYGVPPVYASVGGVLDTVK-----EGVTGFHMGALNPKLDEADADALA	521
Db	652	PCGLNQDIAMSYGVPPVHVGVGGURDTVPPNPFDESGVW-----TFDRAEANKLM	703
QY	522	ATVRASEVFPAGGRYPMVANC-----SQDLSWSKPAQKWGLLEB	563
Db	704	AA-----LWNCILTYDKYKKSWEIGIRGMSQDLSWDAQAQYBEVLVA	747
QY	564	VVY 566	
Db	748	AKY 750	
RESULT 14			
UGS2_ORYSA			
ID	UGS2_ORYSA	STANDARD;	PRT; 626 AA.
AC	Q40739;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Soluble glycogen [starch] synthase, chloroplast precursor		
DE	(EC 2.4.1.11) (SSS).		
OS	Oryza sativa (Rice).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	Ehrhartoideae; Oryzeae; Oryza.		
OX	NCBI_TaxID=4530;		
EN	[1]		
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 114-131.		
RC	STRAIN=cv. Japonica; TISSUE=Seed;		
RX	MEDLINE=94302151; PubMed=7518089;		
RA	Baba T., Nishihara M., Mizuno K., Kawasaki T., Shimada H.,		
RA	Kobayashi E., Ohmishi S., Tanaka K.-I., Arai Y.;		
RT	"Identification, cDNA cloning, and gene expression of soluble starch		
RT	synthase in rice (Oryza sativa L.) immature seeds.";		
RL	Plant Physiol. 103:565-573(1993).		
CC	-! CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-alpha-D-glucosyl)(N) =		
CC	UDP + ((1,4)-alpha-D-glucosyl)(N+1).		
CC	-! PATHWAY: Starch biosynthesis.		
CC	-! SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE.		
CC	-! TISSUE SPECIFICITY: LEAVES AND IMMATURE SEEDS.		
CC	-! MISCELLANEOUS: THREE FORMS OF SOLUBLE STARCH SYNTHASE WERE		
CC	PURIFIED: RSS1, RSS2 AND RSS3.		
CC	-! SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE		
CC	FAMILY.		

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	at the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; D16202; BAA03739.1; -		
DR	InterPro; IPR001296; Glycos_transf_1.		
DR	Pfam; PF00534; Glycos_transf_1; 1.		
KW	Glycogen biosynthesis; Transferrase; Glycosyltransferase;		
KW	transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.		
FT	TRANSIT 1 113		
	CHLOROPLAST (POTENTIAL).		

FT	CHAIN	114	626	SOLUBLE GLYCOCEN [STARCH] SYNTHASE, RSS3.
FT	CHAIN	122	626	SOLUBLE GLYCOCEN [STARCH] SYNTHASE, RSS1.
FT	BINDING	147	147	UDP-GLUCOSE (BY SIMILARITY).
SEQ	SEQUENCE	626 AA;	68451 MW;	03E4182507D26658 CRC64;

Query Match		24.0%;	Score 873.5;	DB 1;	Length 626;
Best Local Similarity		41.5%;	Pred. No. 1.7e-44;		
Matches 197;		Conservative 77;	Mismatches 154;	Indels 47;	Gaps 13;

QY	61	IVMVAEEVAPWSKTTGGLDVTGGPLPIELVKGRHWTIAPRY-----DOYADAWDTSV	111
Db	135	VWFTGASPYAKSGGLDVGCSLPIALALGRHVVMVMPRYMNGALNKNFANAFYTEKH	194
QY	115	VDIM--GE-KVRFHSIKKGVHRVVDIDHPWELAKVWGKTGSKLYGPRSGADYLDNHRF	170
Db	195	IKPCTGGEHEVTFHFHYRSDVWVFVDHFSY-----HRPGNLYGDNFGA-FGDGNFRY	247
QY	171	ALFCKAAIEAARVLPEFG---GEDCVFVANDHWSALYPLVLLKDRYQKPGQFTKAKSVLAI	227
Db	248	TLLCYAAACEAPLILELGGYIYGKCMFVVDHWSALYPLVLLAARYRPGYVRDARSVLVI	307
QY	228	HNIAFGRMGHEAFKDKTLFOAAPDKLAFSDGYAKVYTEATPMEDEKPLTGTGTYKKN	287
Db	308	HNIAHQGEVPASTYPLDGLPPEWYGALEW-----VPEWARRHADKG-----EAVN	354
QY	288	WLKGIITAAADKLVTSPNYATEIAADAAGGVELDTVI--RAKGTEGIVNGMDEEWPKT	345
Db	355	FLGAVVTADRIVTSGVSGWEVTT--AEGGGNELLSSRKSVLNGIDINDDNPST	413
QY	346	DKFLSAPYDONSVAAGKAAAKEALQABLGLPVDPTAPLFAFIGLEQKGVDIILAALPK	405
Db	414	DKELPYHYSVDLL-SGKAKCKAEQLQELGPIRPDVPDLIGIKGLDYQKGLDILTKLAIPD	472
QY	406	ILATPKVQTALGTGKAAAYEKLVAIGTKYKRAKGKGVKFSAPLAHMLTACGAFMLVPSR	465
Db	473	LM-RDNIQTQFVMLGSGDPFGEGWMKSTESGYRDKFRGWMGVFSVPVSHRITAGCDILLMPSR	531
QY	466	FEPCCGLIQLHAMHYGVVAVASTGGGLVDTVK-----EGVTGFHMGALNPDK	511
Db	532	FEPCCNLQIYMGQYGVVPHGHTGGGLDRTVFENFENPAEKGEQGTGWAFSLPTEK	586

```

RESULT 15
UGS2_WHEAT
ID UGS2_WHEAT STANDARD; PRT; 610 AA.
AC Q43654;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Soluble glycoprotein [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11) (Fragment).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OC NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=cv. T.A. Florida; TISSUE=Endosperm;
RA Block M., Loerz H., Luetticke S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -! CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =
CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).
CC -! PATHWAY: Starch biosynthesis.
CC -! SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE (BY
CC SIMILARITY).
CC -! SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U48227; AAB02197.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
FT NON_TER 1 1
FT TRANSIT <1 ? CHLOROPLAST (POTENTIAL).
FT CHAIN ? 610 SOLUBLE GLYCAGEN [STARCH] SYNTHASE.
FT BINDING 7 UDP-GLUCOSE (BY SIMILARITY).
SQ SEQUENCE 610 AA; 67143 MW; 46080A3B7FB87193 CRC64;

```

Query Match	23.94;	Score	869;	DB	1;	Length	610;
Best Local Similarity	35.0%;	Pred.	No. 3e-44;				
Matches	234;	Conservative	94;	Mismatches	222;	Indels	118;
Gaps	22;						
QY	67	EVAPWSKTGGLGDVTCGLPIELYKGRHVMTIAPRY-----DOYADAWTSTVVVDIM--	118				
Db	1	EAAPYAKSGSLGDVCGSLPATAAARGHVMVMVPRYLNGLSSDKNYAKALYAKHILKIPCF	60				
QY	119	--GEKYRPHSHIKKGVHRVWDHPWFLAKVMGKTGSKLYPGSGADYLDNHKRFALFCKA	176				
Db	61	GGSHVETFFHEYRDNDVDWVFVDPH-----SYHRPGS-LYDNTFGA-FGDNQFRYTLLEYA	113				
QY	177	AIFAARVLPGP---GEDCVFVANDHSAIVPVLVKDEYQPKQFTKAKSVLAHNTAFQ	233				
Db	114	ACEAPLLELGGYTYGNCMFVNDWHASIVPVLAAKARYPGVYRDSRSTLVHNLAHQ	173				
QY	234	GRWEEAFKDKLPOAFAFKLAFSDGYAKYYTEATPMEDEKPPLTGKTYKKINWLKGGI	293				
Db	174	GVPEASTYDGLUPPEWYGALW-----VFPWARHRLDKG-----EAVNFKGAV	220				
QY	294	IAADKLVTGSPNYATEATAADAAGVLDVTI--RAKIGIEVNGMDTEENKPTDKFSLA	351				
Db	221	VTADRIVTVSOGYSWEVTT-AEGOGLELLSSRKSVLNGIVNGIDINDWNPDKCLPH	279				
QY	352	PYQNSVYAGKAAKEALQELQELGPVDPDTAPLFAFTGRLEEQKGVDIILAAALPKILATPK	411				
Db	280	HYSVDDIL-SGKACKAKELQELQELGPVREDVPLTGFITGRLDYQKIDLIKWAIPELM-RED	337				
QY	412	VQIAILGTGKAAEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVLVPSRPFCGL	471				
Db	338	VQFVMLGSGDPITFEGWMRSTESTKDKFRGWGVFSVPVSHRITAGCDILLMPSRFPCGL	397				
QY	472	IQLHAMHYGTPVYVASTGGLVDVTVK-----EGVTGFHMGALNPDKLDEADADALAA	522				
Db	398	NQLYAMQYGVVYVHGTTGGDRDVEITFNPFPGKGBEGTGWAFSLPVDKMLWA-----	450				
QY	523	TVRRASEVFAGGRYPFWANCISODLSWSKPAQKQWGLLEVVYVYKGGVAIAKEBKVP	582				
Db	451	-LEIAMSSTFRE-----HKPS--WEGLM-----KRGMTKDH	477				
QY	583	VAEKIPDGLPAVSYAPNTLKPVSASVEGGAAPKVGTTAPAMGAKRATTPSGPSAAAT	642				
Db	478	TWDHAFSSSTRSSGGSWNTPTSC-----RRGLGRSKCSPSALKTSSSS	522				
QY	643	---PKVITYKPAIPATAKPKTAGLK-LAGEASTTSTSENGAASNGNGCASAKTSAAKP	698				
Db	523	FRGPEGYPTLCRPAIVESQACLLWFAG-----SRTYDGA-----AAAVTASGGRQ	570				
QY	699	LYSAATRK	706				
Db	571	LQFWGIRK	578				

Search completed: June 4, 2003, 14:53:07
Job time : 15.5065 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:35:26 ; Search time 35.1742 Seconds
(without alignments)
4147.394 Million cell updates/sec

Title: US-09-980-771A-3

Perfect score: 3641

Sequence: 1 MAVASTRPSARPVINA.....SASKTSRAKPLVSAATKRS 708

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3637	99.9	708	10	O64925	chlamydomon
2	1476.5	40.6	613	10	Q93YB1	pisum sativ
3	1454.5	39.9	604	10	Q9FR03	Q9fr03 perilla fru
4	1441	39.6	607	10	Q9ZS05	Q9zsq5 astragalus
5	1434	39.4	606	10	Q9XIS6	Q9xis6 phaseolus v
6	1428	39.2	608	10	Q93VD9	Q93vd9 ipomoea bat
7	1426	39.2	609	10	Q94LY7	Q94ly7 oryza sativ
8	1419.5	39.0	603	10	Q8SA49	Q8sa49 hordeum vul
9	1418	38.9	609	10	Q8S9C4	Q8s9c4 oryza sativ
10	1409	38.7	605	10	Q9SXK3	Q9sxx3 triticum ae
11	1408	38.7	605	10	Q9FUT6	Q9fu6 triticum ae
12	1405.5	38.6	606	10	Q43012	Q43012 oryza sativ
13	1405.5	38.6	610	10	Q9MAQ0	Q9maq0 arabidopsis
14	1405	38.6	608	10	Q8VYU1	Q8vyul oryza sativ
15	1395.5	38.3	604	10	Q9S7N5	Q9s7n5 triticum ae
16	1394	38.3	605	10	Q9SL88	Q9sls8 triticum tu

17	1394	38.3	605	10	Q9SQ51	Q9sq51 aegilops sp
18	1388.5	38.1	599	10	Q9SQ58	Q9sq58 triticum ae
19	1387.5	38.1	604	10	Q9SL86	Q9sls6 triticum tu
20	1384.5	38.0	605	10	Q9SQ52	Q9sq52 triticum mo
21	1383	38.0	604	10	Q9SL87	Q9sls7 triticum tu
22	1375	37.8	534	10	Q8W2G8	Q8w2g8 triticum ae
23	1374	37.7	574	10	Q9SYU0	Q9syu0 triticum ae
24	1372.5	37.7	604	10	Q9SXK4	Q9sxx4 triticum ae
25	1370.5	37.6	604	10	Q9SL89	Q9sls9 triticum tu
26	1268.5	34.8	565	10	Q9XEN9	Q9xen9 triticum ae
27	900.5	24.7	647	10	Q9LEB9	Q9leb9 triticum ae
28	900.5	24.7	647	10	Q9SQH0	Q9sqh0 aegilops ta
29	900.5	24.7	647	10	Q9SQG9	Q9sgg9 triticum ae
30	899.5	24.7	647	10	Q9LEC0	Q9lec0 triticum ae
31	898	24.7	792	10	Q9MAC8	Q9mac8 arabidopsis
32	895.5	24.6	313	10	Q9LKD3	Q9lkd3 vaquelinia
33	894.5	24.6	643	10	Q9M5A3	Q9m5a3 hordeum vul
34	893.5	24.5	313	10	Q9LKE0	Q9lke0 prunus virg
35	893.5	24.5	313	10	Q9FYU8	Q9fyu8 aruncus dio
36	886.5	24.3	313	10	Q9LKE8	Q9lke8 kageneckia
37	886	24.3	576	10	O64926	chlamydomon
38	885.5	24.3	313	10	Q9LKD1	Q9lkd1 exochorda r
39	882	24.2	630	10	O65365	ipomoea bat
40	879.5	24.2	313	10	Q9LKE5	Q9lke5 amelanchier
41	879.5	24.2	313	10	Q9LKD6	Q9lkd6 vaquelinia
42	878.5	24.1	313	10	Q9LKE7	Q9lke7 kageneckia
43	878.5	24.1	313	10	Q9LKE1	Q9lke1 prinsepia s
44	877.5	24.1	313	10	Q9LKE6	Q9lke6 malus sarge
45	877.5	24.1	313	10	Q9LKD8	Q9lkd8 rosa multif

ALIGNMENTS

RESULT 1

O64925 PRELIMINARY; PRT; 708 AA.
AC O64925.
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Granule-bound starch synthase I precursor (EC 2.4.1.21).
GN STA2.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=137C;
RA D'Hulst C.; Wattedled F., Ral J.-P., Abel G.-J., Kossmann J.,
RA Ball S.G.;
RT "Cloning of a cDNA encoding for the GBSSI in the green alga
RT Chlamydomonas reinhardtii."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Wattedled F., Ball S.G., D'Hulst C.;
RT "Granule-bound starch synthase I: A major enzyme involved in the
RT biogenesis of B-crystallites in starch granules."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF064420; AAC17969.3; -;
DR EMBL; AF433156; AAL28128.1; -;
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR002114; HPr_Serp.site.
DR Pfam; PF00534; Glycos_transf_1.
DR PROSITE; PS00589; PPS_HPR_SER; UNKNOWN_1.
KW Glycosyltransferase; Transferase; Transit peptide.
FT TRANSIT 1 57 POTENTIAL.
FT CHAIN 58 708 GRANULE-BOUND STARCH SYNTHASE I.
SQ SEQUENCE 708 AA; 74623 MW; 7D2A5A07D8606469 CRC64;

Query Match

99.9%; Score 3637; DB 10; Length 708;

```
Best Local Similarity 99.9%; Pred. No. 3.6e-205;
Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAVASTSPSPARPVIVINAAFGYKKTANOLLRELARGSARKSTSRSAVTGATGATCALD 60
D 1 MAVASTSPSPARPVIVINAAFGYKKTANOLLRELARGSARKSTSRSAVTGATGATCALD 60
QY 61 IMVAAEAPKSKTGGGLDVTGGGLPIELVKGRHVRMTIAPRYDQYADAWDTSVVVDIMGE 120
D 61 IMVAAEAPKSKTGGGLDVTGGGLPIELVKGRHVRMTIAPRYDQYADAWDTSVVVDIMGE 120
QY 121 KVRVFSHKGVHRVWIDHPWFLAKVWGKTGSKLYGPRSGADYLDNHRKRALFCKAAIEA 180
D 121 KVRVFSHKGVHRVWIDHPWFLAKVWGKTGSKLYGPRSGADYLDNHRKRALFCKAAIEA 180
QY 181 ARVLPGGEGDCVFFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLAIHNIAPQGRWEEA 240
D 181 ARVLPGGEGDCVFFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLAIHNIAPQGRWEEA 240
QY 241 FKDTKLPOAFAFKLAFSDGYAKVYTFEATPMEDEKPPLTGKTKYKINWLKGGIIAADKLIV 300
D 241 FKDTKLPOAFAFKLAFSDGYAKVYTFEATPMEDEKPPLTGKTKYKINWLKGGIIAADKLIV 300
QY 301 TVSPNYATEIAADAAAGGVELDTVIRAKGIEGIVNGMDIEWNPKDKFELSVPYDONSVA 360
D 301 TVSPNYATEIAADAAAGGVELDTVIRAKGIEGIVNGMDIEWNPKDKFELSVPYDONSVA 360
QY 361 GKAAAKEALQELGLPVPDTPALPFAFIGNLEBQKGVDDIILALPKILATPKVQIAILGTG 420
D 361 GKAAAKEALQELGLPVPDTPALPFAFIGNLEBQKGVDDIILALPKILATPKVQIAILGTG 420
QY 421 KAAEKLVAIGTKYKRAKGVVKSAPLAHMLTAGADFMVPSRFEPCGLLIQHAMHYG 480
D 421 KAAEKLVAIGTKYKRAKGVVKSAPLAHMLTAGADFMVPSRFEPCGLLIQHAMHYG 480
QY 481 TVPVVASTGGLVDTVKEGVTGFHMGALNPDKLDEADALAAATVRRASEVFAGGYPYPMV 540
D 481 TVPVVASTGGLVDTVKEGVTGFHMGALNPDKLDEADALAAATVRRASEVFAGGYPYPMV 540
QY 541 ANCISQDLSWSKPAQKWEGLLEEVYKGGVATAKKEELKVPVAKIPGDLPAVSAPNT 600
D 541 ANCISQDLSWSKPAQKWEGLLEEVYKGGVATAKKEELKVPVAKIPGDLPAVSAPNT 600
QY 601 LKPVASVEGNGAAAPKVTGTAPAMGAWRATTPSGPSAAATPKVTYTKPALPATAKPKT 660
D 601 LKPVASVEGNGAAAPKVTGTAPAMGAWRATTPSGPSAAATPKVTYTKPALPATAKPKT 660
QY 661 AGLKLAGASTTSTSENGAASNGNGASAKTSAAKPLVSAATKSA 708
D 661 AGLKLAGASTTSTSENGAASNGNGASAKTSAAKPLVSAATKSA 708

RESULT 2
Q93YB1 PRELIMINARY; PRT; 613 AA.
AC Q93YB1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Granule-bound starch synthase (EC 2.4.1.11).
GN GBSSI.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RT "Discrete Forms of amylose are synthesised by isoforms of GBSSI in pea."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ345045; CAC69955.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR001917; NHtransf_2.
DR Pfam; PF00534; Glycos_transf_1; 1
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
KW Glycosyltransferase; Transferase; Transf. peptide.
FT TRANSIT 1 85
FT CHAIN 86 613 GRANULE BOUND STARCH SYNTHASE.
SQ SEQUENCE 613 AA; 67626 MW; 5A653461128F97EB CRC64;

Query Match 40.6%; Score 1478.5; DB 10; Length 613;
Best Local Similarity 52.0%; Pred. No. 1.2e-78;
Matches 296; Conservative 75; Mismatches 151; Indels 47; Gaps 11;

QY 39 SARKSTSRSAVTGATGATCALTDIVMAAEVAFWSKTGGGLDVTGGGLPIELVKRGRHVRMTI 98
D 39 SARKSTSRSAVTGATGATCALTDIVMAAEVAFWSKTGGGLDVTGGGLPIELVKRGRHVRMTI 98
QY 67 SKKGVQRKNVPRKGLIIVCGMNLIFVGEVAFWSKTGGGLDVTGGGLPPALSANGHRVMTV 126
D 67 SKKGVQRKNVPRKGLIIVCGMNLIFVGEVAFWSKTGGGLDVTGGGLPPALSANGHRVMTV 126
QY 99 APYDQYADAWDTSVVVDI-----MGEKVFYFHSIKKGVHRVWIDHPWFLAKVWGKTGSKL 154
D 99 APYDQYADAWDTSVVVDI-----MGEKVFYFHSIKKGVHRVWIDHPWFLAKVWGKTGSKL 154
QY 127 TPYDQYKDAWDTNVTIEVKGDRTEKVRFFHCFKRGVDRVDFVHPFIELEKVGKGTGTLK 186
D 127 TPYDQYKDAWDTNVTIEVKGDRTEKVRFFHCFKRGVDRVDFVHPFIELEKVGKGTGTLK 186
QY 155 YGPRSGADYLDNHRKRALFCKAAIEAARVLPF-----GP-GEDEVFVANDWHSALVPV 206
D 155 YGPRSGADYLDNHRKRALFCKAAIEAARVLPF-----GP-GEDEVFVANDWHSALVPV 206
QY 187 YGPAAGDDYQDNLRFSPFQAALEAARVNLKSNKFSGPYGEDVIFVANDWHTALISC 246
D 187 YGPAAGDDYQDNLRFSPFQAALEAARVNLKSNKFSGPYGEDVIFVANDWHTALISC 246
QY 207 LKDEYQPKGQFTKAKSVLAIHNIAPQGRWEEAFKD---TKLPQAAFDKLAFLSDGYAKV 263
D 207 LKDEYQPKGQFTKAKSVLAIHNIAPQGRWEEAFKD---TKLPQAAFDKLAFLSDGYAKV 263
QY 247 YNKMVSQSGIFRNKVVFCIHNIAYQGRF---AFTDYSLLNLPQDKSSDFLDGHVK- 302
D 247 YNKMVSQSGIFRNKVVFCIHNIAYQGRF---AFTDYSLLNLPQDKSSDFLDGHVK- 302
QY 264 YTEATPMEDEKPPLTGKTKYKINWLKGGIIAADKLIVTSPNYATEIAADAAAGGVELDTV 323
D 264 YTEATPMEDEKPPLTGKTKYKINWLKGGIIAADKLIVTSPNYATEIAADAAAGGVELDTV 323
QY 303 -----PIVG---RKINWMKAGIIISHRVLTVPYQAQLVSGDPKGVDELNI 346
D 303 -----PIVG---RKINWMKAGIIISHRVLTVPYQAQLVSGDPKGVDELNI 346
QY 324 IRAKGTGTVNGMDIEWNPKDKFELSVPYDONSVAAGKAAKEALQELGLPVPDTPALP 383
D 324 IRAKGTGTVNGMDIEWNPKDKFELSVPYDONSVAAGKAAKEALQELGLPVPDTPALP 383
QY 347 LRRVGTGTVNGMDVQENPNPSTDKYISIKYDASTVLEGLKALLKEELQAEVGLPVDKNVPL 406
D 347 LRRVGTGTVNGMDVQENPNPSTDKYISIKYDASTVLEGLKALLKEELQAEVGLPVDKNVPL 406
QY 384 FAFIGNLEBQKGVDDIILALPKILATPKVQIAILGTGKAAAYEKLVAIGTKYKRAKGVV 443
D 384 FAFIGNLEBQKGVDDIILALPKILATPKVQIAILGTGKAAAYEKLVAIGTKYKRAKGVV 443
QY 407 IAFIGNLEBQKGVDDIILALPKILATPKVQIAILGTGKAAAYEKLVAIGTKYKRAKGVV 465
D 407 IAFIGNLEBQKGVDDIILALPKILATPKVQIAILGTGKAAAYEKLVAIGTKYKRAKGVV 465
QY 444 KESAPLAHMLTAGADFMVPSRFEPCGLLIQHAMHYGTVPVVASTGGLVDTVKEGVTGFH 503
D 444 KESAPLAHMLTAGADFMVPSRFEPCGLLIQHAMHYGTVPVVASTGGLVDTVKEGVTGFH 503
QY 466 KENVPLAHMIAAGADFILPSRFEPCGLLIQHAMRYGTVPVIVASTGGLVDTVKEGVTGFH 525
D 466 KENVPLAHMIAAGADFILPSRFEPCGLLIQHAMRYGTVPVIVASTGGLVDTVKEGVTGFH 525
QY 504 MGALNP--DKLDEADALAAATVRRASEVFAGGYPYEMVANCISQDLSWSKPAQKWEGLL 561
D 504 MGALNP--DKLDEADALAAATVRRASEVFAGGYPYEMVANCISQDLSWSKPAQKWEGLL 561
QY 526 MGSFNVKCDADVPDVPDYDAIPKTVTKALGVYGSAPAEMLKNCMAQELSNKGPAPKWEVL 585
D 526 MGSFNVKCDADVPDVPDYDAIPKTVTKALGVYGSAPAEMLKNCMAQELSNKGPAPKWEVL 585
QY 562 -----EEVYKGGVATAKKEELKVP 582
D 562 -----EEVYKGGVATAKKEELKVP 582
QY 586 LNLGVPDSEFGIDGQ-EIAPQAKENVATP 613
D 586 LNLGVPDSEFGIDGQ-EIAPQAKENVATP 613

RESULT 3
Q9FR03 PRELIMINARY; PRT; 604 AA.
AC Q9FR03;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Granule-bound starch synthase.
GN GBSSI.
OS Perilla frutescens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; eusterids I; Lamiales; Lamiaceae; Perilla.
OX NCBI_TaxID=48386;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. OKDONG; TISSUE=DEVELOPING SEEDS;
RT Hwang S.-K., Hwang Y.-S.;
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:49:26 ; Search time 12.5962 Seconds
(without alignments)
1653.790 Million cell updates/sec

Title: US-09-980-771A-3
Perfect score: 3641
Sequence: 1 MAVASTSRPSSARPVINA.....SASKTSAKPLVSAATRKSA 708

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pap:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pap:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pap:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pap:*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pap:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1426	39.2	609	3	US-08-941-445A-7	Sequence 7, Appli
2	1377.5	37.8	600	4	US-09-388-743-22	Sequence 22, Appl
3	1358.5	37.3	533	3	US-08-941-445A-5	Sequence 5, Appli
4	1358	37.3	616	4	US-09-388-743-14	Sequence 14, Appl
5	1327.5	36.5	615	4	US-09-388-743-2	Sequence 2, Appli
6	1316.5	36.2	614	4	US-09-388-743-18	Sequence 18, Appl
7	903	24.8	641	4	US-08-836-567-10	Sequence 10, Appl
8	885.5	24.3	649	4	US-09-192-909-2	Sequence 2, Appli
9	884.5	24.3	671	4	US-09-196-390-2	Sequence 2, Appli
10	875.5	24.0	801	4	US-09-388-743-26	Sequence 26, Appl
11	873	24.0	767	4	US-08-836-567-8	Sequence 8, Appli
12	869.5	23.9	583	3	US-08-941-445A-13	Sequence 13, Appl
13	865.5	23.8	539	3	US-08-941-445A-21	Sequence 21, Appl
14	858	23.6	690	4	US-09-388-743-6	Sequence 6, Appli
15	847.5	23.3	698	3	US-08-941-445A-11	Sequence 11, Appl
16	847	23.3	558	4	US-08-836-567-6	Sequence 6, Appli
17	842	23.1	799	4	US-09-196-390-6	Sequence 6, Appli
18	726	19.9	459	4	US-08-836-567-4	Sequence 4, Appli
19	702.5	19.3	669	3	US-08-941-445A-9	Sequence 9, Appli
20	580	15.9	477	1	US-07-735-065-2	Sequence 2, Appli
21	580	15.9	477	1	US-08-469-202-12	Sequence 12, Appl
22	580	15.9	477	2	US-08-484-434C-12	Sequence 12, Appl
23	570.5	15.7	677	4	US-08-836-567-2	Sequence 2, Appli
24	570.5	15.7	1197	4	US-08-836-567-12	Sequence 12, Appl
25	570.5	15.7	1230	2	US-08-968-542C-35	Sequence 35, Appl
26	559	15.4	735	4	US-09-115-704-2	Sequence 2, Appli
27	534.5	14.7	533	4	US-09-388-743-10	Sequence 10, Appl

28 506 13.9 1674 2 US-08-968-542C-12 Sequence 12, Appli
29 240 6.6 64 2 US-08-470-720-15 Sequence 15, Appl
30 215.5 5.9 79 2 US-08-470-720-13 Sequence 13, Appl
31 145.5 4.0 2035 1 US-08-046-585-5 Sequence 5, Appli
32 145.5 4.0 2035 1 US-08-393-703-5 Sequence 5, Appli
33 143.5 4.0 2035 5 PCT-US93-11721-5 Sequence 14, Appl
34 143.5 3.9 1222 2 US-08-470-720-14 Sequence 14, Appl
35 143 3.9 1222 2 US-08-682-517-15 Sequence 15, Appl
36 143 3.9 1252 2 US-08-682-517-9 Sequence 9, Appli
37 132 3.6 520 4 US-09-000-016-7 Sequence 7, Appli
38 132 3.6 520 4 US-09-514-340-7 Sequence 4, Appli
39 132 3.6 734 4 US-09-000-016-4 Sequence 4, Appli
40 132 3.6 734 4 US-09-514-340-4 Sequence 4, Appli
41 132 3.6 823 4 US-09-000-016-2 Sequence 2, Appli
42 132 3.6 823 4 US-09-514-340-2 Sequence 2, Appli
43 130 3.6 1610 4 US-09-513-783A-22 Sequence 22, Appl
44 129 3.5 30 2 US-08-470-720-9 Sequence 9, Appli
45 129 3.5 1125 4 US-09-513-783A-152 Sequence 152, App

ALIGNMENTS

RESULTS
US-08-941-445A-7
; Sequence 7, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P.
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-941-445A-7

Query Match 39.2%; Score 1426; DB 3; Length 609;
Best Local Similarity 48.9%; Pred. No. 2.5e-116;
Matches 302; Conservative 81; Mismatches 172; Indels 62; Gaps 13;
QY 3 VASTSRPSS-----ARPIVINAAPGVKKTAN---OLLRELARGSARKSTSRSA 48
Db 18 IADRSAPSSLLRHFGQLKPRSPAGDATSLSVTTSARATPKQORSVQGRSR---FPVS 74

QY 49 VTGATGATCALTDMVAAEAPVSKTGGTGLDVTGGLPIELVKRGHRYMTIAPRYDQYADA 108
Db 75 VVTAIGA -GNVYVGAEMAPSKTGGTGLDVLGGLPPAMAANGHRVAVISPRDYQYDA 132
QY 109 WDTSVVVDI ----MGEKRVYFHSIKKGVHRVWIDHPWFLAKVWGKTGSKLYGPRSGADYL 164
Db 133 WDTSVVAEIKVADRYVRVFFHCYKRGVDRVFIIDHPSFLEKVGWKTGKIYGPDTGVYDK 192
QY 165 DNHRKRALFCAALEAARVL ----PRGP ----GEDCVFVANDWHSALVPVLLKDEYQPKG 216
Db 193 DNQMRFSILCOALEAPRILNINNPNYFKGTGYGDEVVFCNDWHTGPLASLYKNYQPNQ 252
QY 217 QFTRAKSVLAHNIATFOGRMEEAFKDTKLPOAAFDKLAFLSDGKAVKYTEATPMEDEK 276
Db 253 IYRNVAKVAFCHNSYQGRFAFEDYPELNLSERFSDFIDGY ----DTPVEG ---- 302
QY 277 PLTKYTKKINLWKGIIAADKLVTVSPNYATEIAADAAGVGLDVTIRAKGIEGTNGM 336
Db 303 -----RKINMKAGILEADRVLTSPYAEELISGTARGCELDNIMRLTIGITGVNGM 355
QY 337 DIEWNPKTDFLSAPYDQNSVYAGKAAKAAKEALQAEGLPVDPTAPLFAFICRLLEEOKGV 396
Db 356 DVSEMDPSKDIITAKYDATTALAEKALKEALKEALQAEGLPVDKPIPLAFICRLEEOKGP 415
QY 397 DIILAALPKILATPKVQIAILGTGKAAYEKLVAIGTKYKGRAGVWKFSAPLAHLMTAG 456
Db 416 DVMAAAIPELM -QEDVQIVLIGTGKKFKELLSMEEKYCGKVRVAVKFNAPLAHLINAG 474
QY 457 ADMVLVPSRFPCGLIQLHAMHYGTVPVASTGGLVDVTKGVTGFHMGALNPD--KLDE 514
Db 475 ADVLAVPSRFPCGLIQLQGMRYGTGCACASTGGLVDVTKGVTGFHMGRLSDCKVVEP 534
QY 515 ADADALATVRASVFAAGRYPEMVANCISQDLSWSKPAQKWEGLLEEVYVKGK-- 570
Db 535 SDVKVVAATLKRAIKVGTPTAYEVMVRCNMNODLSWKGPAKKNWNL--LGLGVAGSAPG 592
QY 571 -----VATAKKEIKVP 582
Db 593 IEGDEIAPLAKENVAAP 609
RESULT 2
US-09-388-743-22
; Sequence 22, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; TITLE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Tulipa fosteriana
US-09-388-743-22
Query Match 37.8%; Score 1377.5; DB 4; Length 600;
Best Local Similarity 49.5%; Pred. No. 4.3e-112;
Matches 279; Conservative 79; Mismatches 165; Indels 41; Gaps 10;
QY 42 KTSR--SAVTGATGATCALTDMVAAEAPVSKTGGTGLDVTGGLPIELVKRGHRYMTIA 99
Db 54 KSTPKVGSVVAFTVETGMNLVFTGTGTPSKTGGTGLDVLGGLPPALAAHGRHYMTV 113
QY 100 PRYDQYADAWDTSVVVDI-MGEK--VRYFHSIKKGVHRVWIDHPWFLAKVWGKTGSKLY 155
Db 114 PRYDQYADAWDTNIVLEIKVGDKNMRYVRFHLLHKGVRDVFIDHPWFLAKVWGKTGSKLY 173

QY 156 GPRSGADYLDNHKRFALFCKAAIEARVLPF-----GP-GEDCVFVANDWHSALVPVL 207
Db 174 GPVGTGDDYDQDLNQLRFSLLCOALEAPRVLNLSNNSKYFSGPYGEDVYFTIANDWHTGTLPCY 233
QY 208 LKDEYQPKGQFTKAKSVLAHNIATFOGRMEEAFKDTKLPOAAFDKLAFLSDGKAVKYTEA 267
Db 234 LKSVYKSGLYESAKVAFCIHNMAVQGRFAPDFSLNLPDTOTFKSSDFEFGYTK----- 288
QY 268 TPMEDEKPPITGTYTKKINLWKGIIAADKLVTVSPNYATEIAADAAGVGLDVTIRAK 327
Db 289 -----PVKG-----RKINMKAGILEADIVTVSPYAKELVSGEDRGVGLDNLRLR 336
QY 328 GIEGIVNGMDIEENWPKTDKFLSAPYDQNSVYAGKAAKAAKEALQAEGLPVDPTAPLFAFI 387
Db 337 GVKGVNGMDTNVWNPITDKFTITANYDATMTEAKRVNKQELQAEVGLPVDPIVIVFV 396
QY 388 GELEOKGVDIILAALPKILATPKVQIAILGTGKAAYEKLVAIGTKYKGRAGVWKFSA 447
Db 397 GRLEOKGSDIILAAIPELM-DENVQIILIGTGKHKHLEKELEIEBEQPDKMKRLVAKFN 455
QY 448 PLAHMLTAGADFMVPSRFPCGLIQLHAMHYGTVPVASTGGLVDVTKGVTGFHMGALNPD-- 506
Db 456 PLAHMMAGGDIIPISRFPCGLIQLQGMRYGTGCACASTGGLVDVTKGVTGFHMGAF 515
QY 507 -LNPDKLEADADALAAVVRASVFAAGRYPEMVANCISQDLSWSKPAQKWEGLL---- 561
Db 516 TVECETVDPVDVAGIVTKVRAKLYGTPTAFSEMVCNMAODHWSKPEAKKWEILLGLLE 575
QY 562 ---EYVYVKGKGVATAKKEIKVP 582
Db 576 VDGSEPGDEEIAPLAKENVAAP 599
RESULT 3
US-08-941-445A-5
; Sequence 5, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanning
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 533 amino acids
; TYPE: amino acid

123 RYFHSIRKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKFALECKAAIEAAR 182
157 REFHCYKGVDRVDFVDRHFLFKVWGKTSKLYGPRVTRDQEDNOLRCLLATLETPR 216
183 VL-----PFGP-GEDEVFVANDHWSALVPLVKD-EYQPKGQBTAKSVIAIHNIAFQ 233
217 VLNPNNNKYHSGPKGED-LEFIANDWHTALLPCYLKTIYQAHGIYKNAKVAFCIHNIAYQ 275
234 GRMEEAPKDKLPOAAEDKLAFLSDGYAKVYTEATPMEDEKPLTKGTYKKINWLGKI 293
276 GRFAPEFDSRLNLPDTRFKSSDFIDGYAK-----PIKG---RKINWMKAGI 318
294 IAADKLTVSPNYAIEAADAAGGVDELTVIRAKGIEGVINGMDIEWNPKTDKFLSAPY 353
319 IESDRALIVSPYIAQELVSGIDKGVELNIIRLKTIQGLINGMDINENWNPSTDKITANY 378
354 DONSUYAGAAKAEALQALGELGPDPTAPLFAFTIGRLEEQKGVDIILALPKILATPKVO 413
379 DATTVMEAKPLNKEALQAEVLGPNVKIPVIAFAGLEEQKGSILAEAPKFF-DODVQ 437
414 TAILGTGAAYEKLNAIGTKYKGRAGVVKESAPLAHMLTAGADFMVPSRFEPCGLIQ 473
438 VIVLTGTRKKLQERQALLDEPEPKFRAHMKFNIPLAGIMAGADILVTPSRFEPCGLIQ 497
474 LHAMHYGTVPVVASTGGVLDVTKVGTGFHMG--ALNPDKLDEADADALAAATVRRASEVF 531
498 LOGMYGTPVSMCTITGGVLDVTKVGTGFHMGFSPVECDIADENDVILKIVEAKKRALMY 557
532 AGGRYPPEMVANCISQDLSWSKPAQKWEGLL-----EEVKYVPAEK 586
558 GTPAPEMIONMAQDFSNKGPADKEWKEPILSLGLESGEAGIEGE-EVAPLAKENVATP 615

RESULT 6
US-09-388-743-18
; Sequence 18, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; TITLE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Typha latifolia
US-09-388-743-18

Query Match 36.2%; Score 1316.5; DB 4; Length 614;
Best Local Similarity 47.2%; Pred. No. 9.8e-107;
Matches 271; Conservative 85; Mismatches 165; Indels 53; Gaps 13;

42 KSTRSRAVTCATGAT-----C--ALDIVMVAARVAPWSKTGGIGDVTGGPIELVK 90
60 KTSKQKNGCSTVNLPRPRAVTCGKGMNLVFGAEMAPWSKTGGIGDVLGGLPPALAA 119
91 RGRVMTIAPRYDQADAWDTSVVVDI-MG---EKVYFHSIKKGVHRVWIDHPWFLAKV 146
120 NGRVMVIAFRYDQYMDAWDTDALVELKVGDRCTVREFFHCYKRGVDRVDFVHPMFLAKV 179
147 WGTGSKLYGPRSGADYLDNHRKFALECKAAIEARVLPF-----GP-GEDEVFVAND 198
180 WGTGSKLYGPRSGADYLDNHRKFALECKAAIEARVLPF-----GP-GEDEVFVAND 239
199 WHSALVPLVKDEYQPKGQFTAKSVIAIHNIAFQGRMWEAEAFKDTKLPOAAEDKLAFLSD 258
240 WHTSLPLCYLKSMTVHPRGYKNAKVAFCIHNIAYQGRMWEAEAFKDTKLPOAAEDKLAFLSD 299

259 GYAKVYTEATPMEDEKPLTKGTYKKINWLGKIIADKLTVSPNYATEIAADAAGGV 318
300 GYK-----PVKG---MKINWMAKAGILESDRVFTVSPYTAQELLSGEERG 342
319 ELDTVIRAKGIEGVINGMDIEWNPKTDKFLSAPYDQNSVYAGKAAAEALQALGELPVD 378
343 ELDNILRTVTSITGIVNGMDVNEWNPLTDKYISVNDYAKPVMEAKPLNKEALQAESWLACR 402
379 PTAPLFAFAGLEEQKGVDIILALPKILATPKVQIATLGTGKAAYEKLVAIGTKYKGR 438
403 QSHPCNCHIRLEEQKGSILAASTPEIM-DENVOLIIILGTGKEMENQLESMEEMFPDK 461
439 AGVYKFSAPLAHMLTAGADFMVPSRFEPCGLIQLHAMHYGTVPVVASTGGVLDVTKV 498
462 VRAMKFNAPLAHOMTAGADIIVPSRFEPCGLIQLOQMOYGFPSACSSTGGVLDVTKV 521
499 VTGFHMGALNP--DKLDEADADALAAATVRRASEVFAGGRYPPEMVANCISQDLSWSKPAQK 556
522 KTFHMGPFSAECEVVDLSVKVYTVTKARALKVICTPAFEDMIKMKMAQDLSWKGPKN 581
557 WEGLEEVYVYKGGVATAKK-----EEIKYVPAEK 586
582 WEQVLLNL-----GVAGSEPGYDGEI-VPLAKE 609

RESULT 7
US-08-836-567-10
; Sequence 10, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-836-567-10

```

Query Match      24.8%; Score 903; DB 4; Length 641;
Best Local Similarity 38.1%; Pred. No. 1.7e-70;
Matches 219; Conservative 91; Mismatches 183; Indels 82; Gaps 20;

QY 30 QLLRELARGSARKSTGRSAVTG-----ATGAFCALDIVMVAEVPWPKSTGG 76
Db 89 QLIPSHVAGDATVESHDIIVANDRDILSETERMEETPIKLTNIFVTAAAPYKSTGG 148
QY 77 LGDVTGGLPTELVKGRHVMYTAIPRY-----DQYADAWDTSV--VVDIMG--EKVRYF 125
Db 149 LGDVGSLPMAAARGHRVWVSPRYLNGSPDSDEKYANAVDLVRATVCFEGDAQEAVFY 208
QY 126 HSIKKGVRHWIDHPWFLAKVWKTKSKLYGPRSGADYLDNKHKRFALFCKAAAEARVLP 195
Db 209 HEYRAGVDWVVDHSSYC-----RPGTP-YGDIYGA-FGDNQFRFTLLSHAACEAPLVLP 261
QY 186 FGP--GECVVFVANDHWSALVPVLKDEYQPKQFTKAKSVLAIHNIHAFQGRMWEAEK 242
Db 262 LGGFTYGEKCLFLANDWHAALVPLLAARYPYGVYKARSIVAIIHIAHQVEPAVTYN 321
QY 243 DTKLQAAAFDCKLAFSDGYAKVTEATPMEDEKPPLTGTYKKINLWKGIIAADKLVTY 302
Db 322 NLGLPPQWGAIV--EWIPTWARAHALD-----TGCT--VNVLKGAIAVADRILTV 368
QY 303 SPNYATEIAAAGGVDELDTVI--RAKGTEGIVNGMDIBENPKDKFLSAPYDQNSVYA 360
Db 369 SQGYSWEITT-PEGYGLHELLSSRSQVLNGITNGIDVNDWNPSTDEHTASHYSINDL-S 426
QY 361 GAAAKEALQALQELGVPDPTAPLFAFICRLEBQKVDIILAALPKILATPKVOIALILGTG 420
Db 427 GKVOCKTDLQKELGUPIRPDCEPLIGFICGLDYQKGVDIILSAIPELMQN-DVQVVMGSG 485
QY 421 KAAEKLVAICTYKGRAGKGVKFSAPLAHMLTAGADFMVLPSPRFEPCGLLIQHAMHYG 480
Db 486 EKQYEDWNRHTENLFKDFRAWGVNVPVSHRITAGCDILLMPSRFEPCGLNQLYAMRYG 545
QY 481 TVPVVASTGGVLVDIVK-----EGV---TGFMGALNPKDLDEADADALAAATVRRASEFV 530
Db 546 TIPIVHSTGGLDVTKDNPENYPAQEGIGRGSGTGWTSPLTSEKL---LDTLKLAIV----- 595
QY 531 FAGGRYPE-----MVANCISQDLSWSKPAQKWE 558
Db 596 ---GTYTEHKSSWEGLRMRGMGRDYSWENAAIQYE 627

RESULT 8
us-09-192-909-2
; Sequence 2, Application US/09192909
; Patent No. 6307124
; GENERAL INFORMATION:
; APPLICANT: Jens Kossmann
; APPLICANT: Claus Froberg
; TITLE OF INVENTION: Nucleic acid molecules encoding soluble
; TITLE OF INVENTION: starch synthases from maize
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/192,909
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

Query Match      24.3%; Score 885.5; DB 4; Length 649;
Best Local Similarity 37.8%; Pred. No. 6.1e-69;
Matches 207; Conservative 97; Mismatches 178; Indels 65; Gaps 16;

QY 42 KTSRSVATGATGATCALDIVMVAEVPWPKSTGGGLDVTGGTLPTELKVRGRHVMYTAIPR 101
Db 132 KEQARAKVT-----QNVFVTGEASPYAKSGGLDVCGLSPVALAARGHRVWVMPR 183
QY 102 Y-----DOYADAWDTSVVVDIM--GE-KVRYFHSIKKGVHVRWIDHPWFLAKVWKG 151
Db 184 YLNGTSDKNYANAFYTEKHIRIPCFCGGEHEVTFHEYRDSVDVWVVDHPSY-----HRP 237
QY 152 SKLYPRSGADYLDNKHKRFALFCKAAAEARVLPFP-----GEDCVFVANDHWSALVPVL 208
Db 238 GNLYGDKTGA-FGDNQFRFTLLCYACEAPLVLELGGYIYGNCMFVVDHWSALVPVL 296
QY 209 KDEYQPKQFTKAKSVLAIHNIHAFQGRMWEAEFKDTKLPOAAFDKLAFSDGYAKVYTEAT 268
Db 297 AAKYRPGYGVKDSRSILVHNLAHQGVPEASTYPLGLPPEWYGALEW-----VPEWA 350
QY 269 PMEDEKPPLTGTYKKINLWKGIIAADKLVTVSPNYATEIAAAGGVDELDTVI--RA 326
Db 351 RRHALDKG-----EAVNFKLGAVTADRIVTVSKGYSWEVTT-AEGQGLNELLSRK 402
QY 327 KGIIEVNGMDIEWNPKTDKFLSAPYDQNSVYAKAAKEALQALQELGVLVDPTAPLRAF 386
Db 403 SVLNGIVNGIDINDWNPATIDKICPHYSVDDL-SGAKCKGALQELGILPIRDPVPLIGF 461
QY 387 IGRLEEQGVDIILAALPKILATPKVQITAILGTGKAAVEKLVNAIGTRYKGRAGVVKFS 446
Db 462 IGRLDYQKGLDILIIIPDLN-REDVQFVLMGSDGPELEDNMRSTESIFKOKFRGWGFS 520
QY 447 APLAHLTAGADFMVLPSPRFEPCGLLIQHAMHYGTVPVVASTGGVLVDIVK-----E 497
Db 521 VPVSHRITAGCDILLMPSRFEPCGLNQLYAMQYGVTPVHVHATGLRDTVENFENFGE 580
QY 498 GVTGFHMGALNPKDLDEADADALAAATVRRASEVAGGR--YPMVANCISQDLSWSKPAQ 555
Db 581 OGTGWAFAPLTTN-----MLWTLRTALSTYREHKSSWEGLMKRGMSKDFTDWHAEE 632
QY 556 KWEGGLE 562
Db 633 QYEQIFQ 639

RESULT 9
us-09-196-390-2
; Sequence 2, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; APPLICANT: Lorz, Horst
; APPLICANT: Lutticke, Stephanie

```

```

; APPLICANT: Walter, Lennart
; APPLICANT: Froberg, Claus
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
; TITLE OF INVENTION: SYNTHESIS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 21 588.9
; FILING DATE: 29-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 36 917.7
; FILING DATE: 11-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02793
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: AGREVO-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 671 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-196-390-2

```

```

Query Match 24.3%; Score 884.5; DB 4; Length 671;
Best Local Similarity 34.3%; Pred. No. 7.8e-69;
Matches 248; Conservative 98; Mismatches 245; Indels 133; Gaps 24;

QY 23 GYKTKANOLLRELARGSARKSRSRAVIGATGATCALD-----IVWAAEAVAP 70
DB 9 GYGEAPDLLLE---GIAEDSDISIIAASEQDSEIMDANEOPQAKVTRSVFVTEARAP 65
QY 71 WSKTGLGVDVTGGLPIELVKRGHRYMTIAPRY-----DOYADAMDTSVVVDIM 120
DB 66 YAKSGGLGDCVCSLPALAAARGHRYVMVPRYLNGSSDKNYAKALYIKARHIKIPCFGGSH 125
QY 121 KVRYPHSHKKGVRHWRWIDHPWELAKWVGKTSGLKIPRSGADYLDNHKRFALFCKAAIEA 180
DB 126 EVTFEHEYRNDVWFVDHP-----SYHRPGS-LYGDNFGA-FGDNQFRYTLCLCYAACCA 178
QY 181 ARVLPEGP---GEDCVFVNDHSHALVPVLLKDEYQPGQFQKAKSVLAHNIAFOGRMW 237
DB 179 PLILEGGYIYQNCMFVNDHSHALVPVLLAKYRYPYGYRDSRSTLVHNLHAGOVPE 238
QY 238 EFAFKDTKLPQAFKDLAFSDGAKYVYTEATPMEDEKPLTKTKTKKYNLWKGGLIAD 297
DB 239 ASTYPLDGLPPEWYGALEW-----VPEWARHAKDKG-----EAVNFKGAVVIAD 285
QY 298 KLVTVSPNATETATAADAGGVLDTVI--RAKIEGIVNGMDLEENWPKTKDKFLSAPYDQ 355

```

```

DB 286 RIVTVSQGYSWEVTT-ABGCGQLNELLSRRKSVLNGIVNGIDINDWNPTDKCLPHYSV 344
QY 356 NSVYAGKAAKAEALQAEGLGVPDPTAPFAPFGRLEBKQVDDILLALPKILAPKVOIA 415
DB 345 DDL-SGKAKCAELQELGELGVPREDVPLIGFIRGLDYQKIDLIKNAPELM-REDVQFV 402
QY 416 ILGTGAAAYEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADFMIVPSRFPCLQLH 475
DB 403 MLGSGDP IFEGWMRSTESSYKDKFRGWGVSPVSHRITACDILLMPSRFPCLQLY 462
QY 476 AMHYGTVPVVASTGGGLVDIVK-----EGVTGFHMGALNPDKLDADADALAAVRR 526
DB 463 AMQYGVTVVHGCTGGGLDRIVETNPFCAGKEEGTGWAFSPLTVDKMLWA-----LRT 514
QY 527 ASEVFAGGRYPMEVANCISQDLWSKPAQKWEGLLEEVYKGVGVATAKKEIKVPAEK 586
DB 515 AMSTFER-----HKPS--WESLM-----KRGNTKDHWDH 542
QY 587 IPGDLPAVSYPAPNTLKPVSASVEGNGAAAPKVGTTAPAMGAWRATTSPGSPAAAT--P 643
DB 543 APSSTRSSSGSPWNTPTSC-----RRGLGRSKCESPSALKTSSSRGP 587
QY 644 KVTYKPAKPATAPKPTAGLK-LAGEASTTSTSENGASNGNGASASKTSAAKPLVSA 702
DB 588 EGYPTLCRPATVESQCACLLWFAG-----SRIYDGA-----AAAVTASGGRLQFW 635
QY 703 ATRK 706
DB 636 GIRK 639

RESULT 10
US-09-388-743-26
; Sequence 26, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; FILE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Typha latifolia
; US-09-388-743-26

Query Match 24.0%; Score 875.5; DB 4; Length 801;
Best Local Similarity 36.5%; Pred. No. 6.4e-68;
Matches 225; Conservative 85; Mismatches 199; Indels 107; Gaps 19;

QY 2 AVASTSRPSS-----ARPIVINAASFQVKKTANOLLRELARGSARKSTRS----- 47
DB 156 AVDAFGRPPEPSLGTITKILSPFYLEASDCAKENAEADLVEAKLDSVHVKKDNLNPGEE 215
QY 48 ---AVTGATGATCALDIVMVAEVAEPWSTGGIGDVTGGLPIELVKGHRVMTIAPRYDQ 104
DB 216 VPLPLAGAN---VMNLIIVAAECAPWSKTGGIGDVGAGALPKALARHVRVMPVARYGN 271
QY 105 YADAMDTSV--VVDIMGE--KVRYPHSHKKGVRHWRWIDHPWELAKWVGKTSGLK-GRPS 159
DB 272 YAEFQDVGVRKYTKYKVGQDMVEVTFHAYIDGVDFVMDSPDFRHR-----GNRIYEGNR- 325
QY 160 GADYLDNHKRFALFCKAAIEAARVLP-----FGPEGDCVFVANDHSHALVPVLLKDEYQ 214
DB 326 ---VDILKMLIFCKAAVEVPWHPVPCGFCYGDG-NLAFTINDWHTALLPVLYKAYRD 380
QY 215 KGOFTKAKSVLAHTNTAQGRMMFEAFKDKLPQAFKDLAFSDGAKYVYTEATPMEEDE 274

```

```

Db 381 NGLMKYARSLVITHNIAHQGRGPVDFKFEVGLPDHYLDLFLRYD----- 424
QY 275 KPPLTGTXYKINLWKGIIAADKLVTYSPNYATEIAADAAAGGVVELDTVIRAKG--IEGI 332
Db 425 --PVGG--EHLNIFAAGLTADRVTVSHGYAWELKT-SEGWWGLHEIILNESNWKFGI 478
QY 333 VNGMDIEWPKDKFLSAP-----YDONSUYAGKAAAEALQAEGLPVPDPTAPLEAFIG 388
Db 479 VNGIDAKESPEFDVHLKSGDYNISLDTLEMGKPCVKCAALQREVGLPVRDNPVILAFIG 538
QY 389 RLBEOKGVDIILALPKILATPKVQIAILGTGKAAAEKLYNAIGTKYKGRAGKGVKFSAP 448
Db 539 RLHQKGVDIILAEAPWIV-SHDVQVVMGLGTGRODLENNLRNEEGHRDKVRAWAFSVK 597
QY 449 LAHMLTAGADFMVPSRFEPCGLIQLHAMHYGTVPVYASTGGVLVDIVKEGVTGHEHMGALN 508
Db 598 MAHRIITAGADILMPSRFEPCGLNQLYAMMIGTIPVVHVGGLRDTVTO----- 646
QY 509 PKLDEADADALATVRRASEVFAGGRYPEMVANCIS-----QDLSW 550
Db 647 ---FDPFNESSGLGWTFDRAE---AGKLIHALNNCLNTYWNKYDSWKGLQTRGMQDLSW 699
QY 551 SKPAQKWEGLLEEVY 566
Db 700 DNAAQOYEDVLVAKY 715

RESULT 11
US-08-836-567-8
; Sequence 8, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836.567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 767 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: protein
US-08-836-567-8

```

```

Query Match 24.0%; Score 873; DB 4; Length 767;
Best Local Similarity 39.7%; Pred. No. 9.9e-68;
Matches 211; Conservative 78; Mismatches 183; Indels 60; Gaps 14;

```

```

QY 56 TCALDIVMAAEVAPWSKTGGLGVDTGTELKRGHRVMTIAPRYDQYADADWTSV-- 113
Db 273 TNVNNIIVASECAPWSKTGGLGVAGALPKALARRGHRVMAVAPRYDVPQDSGVYRK 332
QY 114 --VVDIMGEKVRYFHSIKKGVRHVIDHPFLAKVWGTGSKLYCPRSGADYLDNHRFA 171
Db 333 IYKVGDDGVETTFQAFIDGVDFID-----SHMERHIGNNIY----GGNRVDILKRW 383
QY 172 LFCKAAATEAARVLP-----FGPGEDCVFVANDHSAVLPVLLKDEYQPKGQFTAKSVLA 226
Db 384 LFCKAAIEVPHVPCGGVGYGDG--NLVFIANDHNTALLPVYLKAYYRDNGIMNTRSVLV 442
QY 227 IHNIAFGRMWEAFKDKLPQAAAFDKLAFSDGYAKVYTEATPMEDEKPLTGKTYKKI 286
Db 443 IHNIAHQGRGLEDFSYVDLPPHYMDPKLYD-----PVGGEHF--- 481
QY 287 NWLKGIIAADKLVTYSPNYATEIAADAAAGVELDTVIRAKG--IEGVNGMDIEWNPK 344
Db 482 NIFAAGLKTADRVTVSHGYSWELKT--SQGGWGLHQIINENDWKLOGIVNGIDTKEWNP 540
QY 345 TDKFLSAP-----YDONSUYAGKAAAEALQAEGLPVPDPTAPLEAFIGRLEOKGVDIIL 400
Db 541 LDVHLQSDGYMNYSLDTLQTKPCKAAALQKELGSLPVRDVPVLTGTRGLDPQKGVDLIA 600
QY 401 AALPKILATPKVQIAILGTGKAAAEKLYNAIGTKYKGRAGKGVVYKFSAPLAHMLTAGADM 460
Db 601 EASAWMMG--QDVQLVMLGTGRDRLEQMLROFECQHNNDKINGWVGFVSKTSHRITAGADIL 659
QY 461 LVPSRPEPCGLIQLHAMHYGTVPVYASTGGVLDTVK-----EGVTGFHMGALNPKLDE 514
Db 660 LMPSRPEPCGLNQLYAMKYGTIPVVHVGGLRDTVQFPDFNBSGLGWTFSRAEASQLIH 719
QY 515 ADADALAATVRRASEVFAGGRYPEMVANCISQDLSWSKPAQKWEGLLEEVY 566
Db 720 ALGNCL-LTYREYKSWEG-----IQTRCMTQDLSWDNAAQNYEEVLIAKY 765

```

```

RESULT 12
US-08-941-445A-13
; Sequence 13, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanning
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941.445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:

```

COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/941.445A
FILING DATE: 30-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,855
FILING DATE: 30-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Winner, Ellen P
REGISTRATION NUMBER: 28,547
REFERENCE/DOCKET NUMBER: 89-97
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-941-445A-21

NAME: Winner, Ellen P
REGISTRATION NUMBER: 28,547
REFERENCE/DOCKET NUMBER: 89-97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-941-445A-13

Query Match 23.8%; Score 865.5; DB 3; Length 539;
Best Local Similarity 40.0%; Pred. No. 2.6e-67;
Matches 198; Conservative 84; Mismatches 158; Indels 55; Gaps 14;
QY 42 KSTSAVIGATGATCAGTCAIDIVVAAEVAPSWKGGTGLDVTGGLPIELVKRGHVTIAPR 101
DB 40 KEQARAKVTQS-----IVFVTGEASPYAKSGGLDVGCSLPVLAARAGHRVVMVMPR 91
QY 102 Y-----DQYADAWTSSVVVDIM--GE-KVRYFHSIKKGVHRVMDHPWFLAKVWGKTG 151
DB 92 YLNGTSDKNYANAFYTEKHIRIPCFGEHEVIFFFHEYRSDVWVFDHPSY-----HRP 145
QY 152 SKLYGPRGADYLDNHRKFAFCFAAIEARVLPFGP---GEDCVFVANDWHSALVPVLL 208
DB 146 GNLGDKFGA-FGDNQFRYTLCCYAACEAPLILELGGYIYGQNCMFVNDWHSALVPVLL 204
QY 209 KBYQKPGOPTAKSVLAIHNTAFQGRMWEAEFKDKLPQAAFDKLAFAFDGKAYVYTEAT 268
DB 205 AAKYRPGYVKDSRSILVHNLAHQGVPEASTYPDGLGPPPEWYGALEW-----VPEWA 258
QY 269 PEEDEKPPLTGKTYKKINLKGIIAADKLVTSPNYATEIAADAAGGVDELDTVI--RA 326
DB 259 RRHALDKG-----EAVNFKGAVVTADRIYTVSKYSWEVTT-AEGGQGLNELLSRK 310
QY 327 KGIEGVNGMDIEWNPETDKFLSAPYDQNSVYAGKAAKEALQAEGLGLPVDPTAPLFAF 386
DB 311 SVLVINGVNDINDWNPATDKCIPCHYSVDLL-SGAKCKGALQKELGLTRPDVPLIGF 369
QY 387 IGRLEQKGVDTIILAAALPKILATPKVOIALGTGKAAAYEKLVAIGTKYKGRAGVYKFS 446
DB 370 IGRLDYQKGDILLIQLIIPDLM-REDVQFVMLGSGDPELDWNRSTESIFKDKFRGWGFS 428
QY 447 APLAHLMTAGADFMVPSRFEPCGLIOLHAMHYGTVPVASTGTGLVDTVK-----E 497
DB 429 VPVSHRITAGCDILLMPSRFEPCGLNQLYAMQYGTVPVHATGGLRDTVENFNGFGE 488
QY 498 GVTGFHMGALNPDKL 512
DB 489 QGTGWAFAPLTENN 503

Query Match 23.9%; Score 869.5; DB 3; Length 583;
Best Local Similarity 38.3%; Pred. No. 1.3e-67;
Matches 209; Conservative 90; Mismatches 174; Indels 73; Gaps 17;
QY 6 TSFSSARPIVINA-----ASFQVYKKT---ANQLRELARGSARKSTSRSAVT 50
DB 36 TGEFASPTPPPDAGLDGLLEPEGIAEGSIDNTIVVASEQDSEIVVG---KEQARAKVT 92
QY 51 GATGATCAGTCAIDIVVAAEVAPSWKGGTGLDVTGGLPIELVKRGHVTIAPR-----DQ 104
DB 93 QS-----IVFVTGEASPYAKSGGLDVGCSLPVLAARAGHRVVMVMPRYLNGTSDKN 144
QY 105 YADAWTSSVVVDIM--GE-KVRYFHSIKKGVHRVMDHPWFLAKVWGKTGSKLYGPRSG 160
DB 145 YANAFYTEKHIRIPCFGEHEVIFFFHEYRSDVWVFDHPSY-----HRPGLNGDKFG 198
QY 161 ADYLDNHRKFAFCFAAIEARVLPFGP---GEDCVFVANDWHSALVPVLLKDEYQPKGQ 217
DB 199 A-FGDNQFRYTLCCYAACEAPLILELGGYIYGQNCMFVNDWHSALVPVLLAAKYRPGV 257
QY 218 FTKAKSVLAIHNTAFQGRMWEAEFKDKLPQAAFDKLAFAFDGKAYVYTEATPEEDEKPP 277
DB 258 YKDSRSILVHNLAHQGVPEASTYPDGLGPPPEWYGALEW-----VPEWARHALDKG- 310
QY 278 LNKTYKKINLKGIIAADKLVTSPNYATEIAADAAGGVDELDTVI--RAKGIEGVNG 335
DB 311 -----EAVNFKGAVVTADRIYTVSKYSWEVTT-AEGGQGLNELLSRKSVLNGVNG 363
QY 336 MDIEWNPETDKFLSAPYDQNSVYAGKAAKEALQAEGLGLPVDPTAPLFAFTRLEEQK 395
DB 364 IDINDWNPATDKCIPCHYSVDLL-SGAKCKGALQKELGLTRPDVPLIGFTRLDYQK 422
QY 396 VDILILAAALPKILATPKVOIALGTGKAAAYEKLVAIGTKYKGRAGVYKFSAPLAHMLTA 455
DB 423 IDLIQLIIPDLM-REDVQFVMLGSGDPELDWNRSTESIFKDKFRGWGFSVPVSHRITA 481
QY 456 GADFMVPSRFEPCGLIOLHAMHYGTVPVASTGTGLVDTVK-----EGVTFHMGGA 506
DB 482 GCDILLMPSRFEPCGLNQLYAMQYGTVPVHATGGLRDTVENFNGFGEQGTGWAFAP 541
QY 507 LNPDKL 512
DB 542 LITENN 547

Query Match 23.8%; Score 865.5; DB 3; Length 539;
Best Local Similarity 40.0%; Pred. No. 2.6e-67;
Matches 198; Conservative 84; Mismatches 158; Indels 55; Gaps 14;
QY 42 KSTSAVIGATGATCAGTCAIDIVVAAEVAPSWKGGTGLDVTGGLPIELVKRGHVTIAPR 101
DB 40 KEQARAKVTQS-----IVFVTGEASPYAKSGGLDVGCSLPVLAARAGHRVVMVMPR 91
QY 102 Y-----DQYADAWTSSVVVDIM--GE-KVRYFHSIKKGVHRVMDHPWFLAKVWGKTG 151
DB 92 YLNGTSDKNYANAFYTEKHIRIPCFGEHEVIFFFHEYRSDVWVFDHPSY-----HRP 145
QY 152 SKLYGPRGADYLDNHRKFAFCFAAIEARVLPFGP---GEDCVFVANDWHSALVPVLL 208
DB 146 GNLGDKFGA-FGDNQFRYTLCCYAACEAPLILELGGYIYGQNCMFVNDWHSALVPVLL 204
QY 209 KBYQKPGOPTAKSVLAIHNTAFQGRMWEAEFKDKLPQAAFDKLAFAFDGKAYVYTEAT 268
DB 205 AAKYRPGYVKDSRSILVHNLAHQGVPEASTYPDGLGPPPEWYGALEW-----VPEWA 258
QY 269 PEEDEKPPLTGKTYKKINLKGIIAADKLVTSPNYATEIAADAAGGVDELDTVI--RA 326
DB 259 RRHALDKG-----EAVNFKGAVVTADRIYTVSKYSWEVTT-AEGGQGLNELLSRK 310
QY 327 KGIEGVNGMDIEWNPETDKFLSAPYDQNSVYAGKAAKEALQAEGLGLPVDPTAPLFAF 386
DB 311 SVLVINGVNDINDWNPATDKCIPCHYSVDLL-SGAKCKGALQKELGLTRPDVPLIGF 369
QY 387 IGRLEQKGVDTIILAAALPKILATPKVOIALGTGKAAAYEKLVAIGTKYKGRAGVYKFS 446
DB 370 IGRLDYQKGDILLIQLIIPDLM-REDVQFVMLGSGDPELDWNRSTESIFKDKFRGWGFS 428
QY 447 APLAHLMTAGADFMVPSRFEPCGLIOLHAMHYGTVPVASTGTGLVDTVK-----E 497
DB 429 VPVSHRITAGCDILLMPSRFEPCGLNQLYAMQYGTVPVHATGGLRDTVENFNGFGE 488
QY 498 GVTGFHMGALNPDKL 512
DB 489 QGTGWAFAPLTENN 503

Query Match 23.9%; Score 869.5; DB 3; Length 583;
Best Local Similarity 38.3%; Pred. No. 1.3e-67;
Matches 209; Conservative 90; Mismatches 174; Indels 73; Gaps 17;
QY 6 TSFSSARPIVINA-----ASFQVYKKT---ANQLRELARGSARKSTSRSAVT 50
DB 36 TGEFASPTPPPDAGLDGLLEPEGIAEGSIDNTIVVASEQDSEIVVG---KEQARAKVT 92
QY 51 GATGATCAGTCAIDIVVAAEVAPSWKGGTGLDVTGGLPIELVKRGHVTIAPR-----DQ 104
DB 93 QS-----IVFVTGEASPYAKSGGLDVGCSLPVLAARAGHRVVMVMPRYLNGTSDKN 144
QY 105 YADAWTSSVVVDIM--GE-KVRYFHSIKKGVHRVMDHPWFLAKVWGKTGSKLYGPRSG 160
DB 145 YANAFYTEKHIRIPCFGEHEVIFFFHEYRSDVWVFDHPSY-----HRPGLNGDKFG 198
QY 161 ADYLDNHRKFAFCFAAIEARVLPFGP---GEDCVFVANDWHSALVPVLLKDEYQPKGQ 217
DB 199 A-FGDNQFRYTLCCYAACEAPLILELGGYIYGQNCMFVNDWHSALVPVLLAAKYRPGV 257
QY 218 FTKAKSVLAIHNTAFQGRMWEAEFKDKLPQAAFDKLAFAFDGKAYVYTEATPEEDEKPP 277
DB 258 YKDSRSILVHNLAHQGVPEASTYPDGLGPPPEWYGALEW-----VPEWARHALDKG- 310
QY 278 LNKTYKKINLKGIIAADKLVTSPNYATEIAADAAGGVDELDTVI--RAKGIEGVNG 335
DB 311 -----EAVNFKGAVVTADRIYTVSKYSWEVTT-AEGGQGLNELLSRKSVLNGVNG 363
QY 336 MDIEWNPETDKFLSAPYDQNSVYAGKAAKEALQAEGLGLPVDPTAPLFAFTRLEEQK 395
DB 364 IDINDWNPATDKCIPCHYSVDLL-SGAKCKGALQKELGLTRPDVPLIGFTRLDYQK 422
QY 396 VDILILAAALPKILATPKVOIALGTGKAAAYEKLVAIGTKYKGRAGVYKFSAPLAHMLTA 455
DB 423 IDLIQLIIPDLM-REDVQFVMLGSGDPELDWNRSTESIFKDKFRGWGFSVPVSHRITA 481
QY 456 GADFMVPSRFEPCGLIOLHAMHYGTVPVASTGTGLVDTVK-----EGVTFHMGGA 506
DB 482 GCDILLMPSRFEPCGLNQLYAMQYGTVPVHATGGLRDTVENFNGFGEQGTGWAFAP 541
QY 507 LNPDKL 512
DB 542 LITENN 547

RESULT 14
US-09-388-743-6
; Sequence 6, Application US/09388743

RESULT 13
US-08-941-445A-21
; Sequence 21, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO

us-09-980-771a-3.ra1

Tue Jun 10 09:03:09 2003

Db 695 KX 696

Search completed: June 4, 2003, 14:58:30
Job time : 18.5962 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:56:21 ; Search time 21.8651 Seconds
(without alignments)
3342.965 Million cell updates/sec

Title: US-09-980-771a-3

Perfect score: 3641

Sequence: 1 MAVASTSRSSARPVINA.....SASKTSAAKPLVSAATRKSA 708

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB_PEP.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB_PEP.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB_PEP.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB_PEP.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB_PEP.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB_PEP.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB_PEP.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB_PEP.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB_PEP.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB_PEP.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB_PEP.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB_PEP.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB_PEP.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB_PEP.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1467.5	40.3	636	9 US-10-138-075-4	Sequence 4, Appli
2	1449.5	39.8	609	9 US-10-138-075-2	Sequence 2, Appli
3	1388.5	38.1	599	9 US-10-138-075-5	Sequence 5, Appli
4	884.5	24.3	671	10 US-09-952-677-2	Sequence 2, Appli
5	842	23.1	799	10 US-09-952-677-6	Sequence 6, Appli
6	562.5	15.4	915	9 US-10-163-214-6	Sequence 6, Appli
7	559	15.4	909	9 US-10-163-214-2	Sequence 12, Appli
8	555.5	15.3	914	9 US-10-163-214-12	Sequence 2, Appli
9	547.5	15.0	874	9 US-10-163-214-13	Sequence 13, Appli
10	410.5	11.3	474	12 US-10-007-693-99	Sequence 99, Appli
11	402.5	11.1	476	12 US-10-007-693-69	Sequence 69, Appli
12	385.5	10.6	293	9 US-10-163-214-10	Sequence 10, Appli
13	215.5	5.9	117	10 US-09-739-438-2	Sequence 2, Appli
14	206	5.7	409	9 US-09-738-626-4738	Sequence 4738, Ap
15	189	5.2	381	9 US-09-738-626-5896	Sequence 5896, Ap
16	164	4.5	143	10 US-09-739-438-4	Sequence 4, Appli
17	155.5	4.3	191	10 US-09-924-358-22	Sequence 22, Appli
18	155.5	4.3	418	9 US-09-738-626-3951	Sequence 3951, Ap
19	143	3.9	1222	10 US-09-137-531-15	Sequence 15, Appli

20	143	3.9	1252	10 US-09-137-531-9	Sequence 9, Appli
21	136.5	3.7	385	10 US-09-767-041-18	Sequence 18, Appli
22	135	3.7	2993	9 US-09-738-626-6239	Sequence 6239, Ap
23	132	3.6	2910	9 US-10-124-800-2	Sequence 2, Appli
24	130	3.6	1610	9 US-10-100-957A-22	Sequence 22, Appli
25	129	3.5	1125	9 US-10-100-957A-152	Sequence 152, Appli
26	128	3.5	3241	10 US-09-841-786-1	Sequence 1, Appli
27	126	3.5	387	9 US-09-738-626-3890	Sequence 3890, Ap
28	125.5	3.4	398	10 US-09-934-899-16	Sequence 16, Appli
29	125.5	3.4	398	10 US-09-934-868-36	Sequence 36, Appli
30	124.5	3.4	1704	9 US-09-991-262-40	Sequence 40, Appli
31	124	3.4	379	9 US-09-738-626-6655	Sequence 6655, Ap
32	124	3.4	1461	9 US-10-021-955-88	Sequence 88, Appli
33	123.5	3.4	413	9 US-09-738-626-4025	Sequence 4025, Ap
34	123	3.4	1528	9 US-09-945-917-3	Sequence 3, Appli
35	123	3.4	1583	9 US-09-945-917-4	Sequence 4, Appli
36	122.5	3.4	2756	9 US-10-331-061-7	Sequence 7, Appli
37	122.5	3.4	3816	9 US-09-808-880-3	Sequence 3, Appli
38	121	3.3	339	9 US-09-847-208-116	Sequence 116, Appli
39	121	3.3	636	9 US-09-738-626-5590	Sequence 5590, Ap
40	120	3.3	2478	10 US-09-815-242-5816	Sequence 5816, Ap
41	120	3.3	2478	10 US-09-815-242-12967	Sequence 12967, A
42	118.5	3.3	372	10 US-09-815-242-5277	Sequence 5277, Ap
43	118.5	3.3	384	10 US-09-815-242-12605	Sequence 12605, A
44	118.5	3.3	795	9 US-09-738-626-3810	Sequence 3810, Ap
45	118	3.2	267	10 US-09-852-555-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-138-075-4
; Sequence 4, Application US/10138075
; Publication No. US20030087369A1
; GENERAL INFORMATION:
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Orozco, Emil M.
; TITLE OF INVENTION: Granule-Bound Starch Synthase
; FILE REFERENCE: BB1474 NA
; CURRENT APPLICATION NUMBER: US/10138,075
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/288,315
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Glycine max
US-10-138-075-4

QY	24	VKTNQLRLRLARSARKSTSRSAVTGATGATCALDIVMVAEAPWSKTGGLGDTGG	83
DB	79	IKTNVAQMRKPGQKNRPKGM-----ITCGMTFIIIGTEVAPWCCKTGLGDLVGG	130
QY	84	LPFELVKRGHRVMTAPRYDOYADAWDSVVDI-----MGEKRVFHSKKGVHRVIDH	139
DB	131	LPALAGTCHRVMTIVPRYDQYKDAWDSVLEVKGDRTEKRVFFHCYKRGVDRVVDH	190
QY	140	PWFLAKVWGKTKLYGPRSGADYLDNKHFKALFCKAATEAARVLPF-----GP-GED	191
DB	191	PWFLEKVGKTKQKLYGPTGNDYEDNQRLSFQQAALAPRVLSLSSKSYFGPYGED	250
QY	192	CVFVANDHSLVPLVLLKDEYQPKQFTKAKSVLAHNIAFOGRMWEBAFKDKLPQAAF	251

Db 251 VIFVANDWHTALIPCYLKSMTQSRGIYTNARVVFCHINAYOGRFAFADFSLNLPDOFK 310
QY 252 DKLAISDGYAKYITTEATPMEDEKXPLGKTKYKKNLWKGIIAADKLVTSVPNATETA 311
Db 311 SSFDFIDGHVK-----PVVG---RKNLWKRAGLIESFVITVSPNAYAKELV 353
QY 312 ADAAGVELDVTIRA-----KGIEGIVNGMDIEWNPKTDKFLSAPYDONSIVYAGKAAAKE 367
Db 354 SGPDGKVELDNIIRKIDDDGRLVGIVNGMDVOEWNPETDKYIAVKYDVSTVLEAKALLKE 413
QY 368 ALQAEGLPDPVDTAPLAFIAGRLIEQKGVDIILALPKLILATPKVQIAILGTGKAAAYEKL 427
Db 414 ALQAEVGLPVDNRNIPILIGFIRLEQKGSIIAAEALPOFI-KONVOLVALGTGKKOMEQ 472
QY 428 VNAIGTKYKRAKGVVVKFSAPLAHMLTAGADFMVPSREPCGLIOLHAMHYGTVPVVAS 487
Db 473 LEELETSYDPKARGVAKFNVPLAHMIIAGADFILVPSREPCGLIOLQAMRYGSVPIVAS 532
QY 488 TGLVDTVKREGVTFHMGALNP--DKLDEADALAAATVRRASEVFAGGRYPPEMVANCIS 545
Db 533 TGLVDTVKREGVTFHMGALNP--DKLDEADALAAATVRRASEVFAGGRYPPEMVANCIS 545
QY 546 QDLSWKPAAKKEGLEEVY-----GKGG--VATAKKEIKVP 582
Db 593 QDLSWKPAAKKEGLEEVY-----GKGG--VATAKKEIKVP 582

RESULT 2
US-10-138-075-2
; Sequence 2, Application US/10138075
; Publication No. US20030087369A1
; GENERAL INFORMATION:
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Orozco, Emil M.
; TITLE OF INVENTION: Granule-Bound Starch Synthase
; FILE REFERENCE: BB1474 NA
; CURRENT APPLICATION NUMBER: US/10/138,075
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/288,315
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Zea mays
US-10-138-075-2

Query Match 39.8%; Score 1449.5; DB 9; Length 609;
Best Local Similarity 48.7%; Pred. No. 2.4e-94;
Matches 302; Conservative 85; Mismatches 164; Indels 69; Gaps 14;
QY 6 TSFSPARPVINAASEGVKKTANQLRELARGS-----ARKSTSRSAVTGATG- 54
Db 16 TNRESALK---QAPHMOFOCCNGGLRFLSKHSQSRSKIQVAKRRATDNGIHPKTTGH 71
QY 55 -----ATCAlDlVWAAEVAPWSKTGGLGDVTGGLPIELVKRGHRVMTIAPRYDOYADAW 109
Db 72 RAPIVCSAGMTIVFAVEVHPWCKTGGLGDVVGGLPPALAAAMGRVMTIAPRYDOYKDAW 131
QY 110 DTSVVVDI-MG---EKRYFHSIKKGVHRVMDHPFLAKVWGTGSKLYGPRSGADYLD 165
Db 132 DTSVLVEVNGDVTETVREFCYKRGVDRVDFVHPMFLKVMWGTGAKLYGPTGTDYRD 191
QY 166 NKRFPALCKAAEAARVLFP-----GP-GEDCVFVANDWHSALVPLVKDEYQPKGQ 217
Db 192 NOLRCELCLAAEAAPVLNENNSYFSGPYGDFVDFVANDWHTAILPCYLKSMYKPKNGI 251
QY 218 FTKAKSVLAHNIATQGRWEEAFKDTKLPOAFDKLAFSDGYAKVYTEATPMEDEKPP 277

Db 252 YNNKAVAFCHINAYOGRFAFADFSLNLPDSFLPSDFIDGHVK-----P 297
QY 278 LTGKTKYKKNLWKGIIAADKLVTSVPNATETAADAAGVELDVTIRAKGIB-GIVNGM 336
Db 298 VLG---RKNLWKRAGLIESDLVTSVPNAYAKELVSGDVGELDKPLKPLGIVNGM 354
QY 337 DIEWNPKTDKFLSAPYDONSIVYAGKAAAKEALQAEGLPDPVDTAPLAFIAGRLIEQKGV 396
Db 355 DVEWNPSTDKYISAKYDATVTFEARLNKERLOAEVGLPVDSSIPVIVFVGRLEEQKGS 414
QY 397 DIILALPKLILATPKVQIAILGTGKAAAYEKLVAIGTKYKRAKGVVVKFSAPLAHMLTAG 456
Db 415 DIILALPKLILATPKVQIAILGTGKAAAYEKLVAIGTKYKRAKGVVVKFSAPLAHMLTAG 456
QY 457 ADEMLVPSREPCGLIOLHAMHYGTVPVASTGGLVDVTGKGVTFHMGALNP--DKLDE 514
Db 474 ADLIVPSREPCGLIOLQAMRYGVPICTSGTGLVDTVBEGVTFHMGFSNFVCEVDP 533
QY 515 ADALAAATVRRASEVFAGGRYPPEMVANCISQDLSWKPAAKKEGLEEVYVYKGG----- 569
Db 534 ADTAVASTVTRALKQYDTPAFHEMVONCMAQDLSWKPAAKKEGLEEVYVYKGG----- 569
QY 570 -----GVATAKKEIKVP 582
Db 590 AGIDDAEIEIAPLAKENVATP 609

RESULT 3
US-10-138-075-5
; Sequence 5, Application US/10138075
; Publication No. US20030087369A1
; GENERAL INFORMATION:
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Orozco, Emil M.
; TITLE OF INVENTION: Granule-Bound Starch Synthase
; FILE REFERENCE: BB1474 NA
; CURRENT APPLICATION NUMBER: US/10/138,075
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/288,315
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-138-075-5

Query Match 38.1%; Score 1388.5; DB 9; Length 599;
Best Local Similarity 48.5%; Pred. No. 4.8e-90;
Matches 288; Conservative 84; Mismatches 167; Indels 55; Gaps 13;
QY 18 NAASEGVKKTANQLRELARGS-----GSARKSTSRSA-VTGATGATCALDIVWAAEVAPWSKT 74
Db 32 NEVMFLSMRNTKTLAKRRATNYGTHRNSSRTPAPIVCSGTG---MPIIFATVETVHPWCKT 87
QY 75 GGLGDVTGGLPIELVKRGHRVMTIAPRYDOYADAWDTSVVVDIM-----GEKRYFHSIKK 130
Db 88 GGLGDVVGGLPPALAAAMGRVMTIAPRYDQKDTWDTNVLVEVIVGDRTEVTFVRFPHCYKR 147
QY 131 GVHRVMDHPFLAKVWGTGSKLYGPRSGADYLDNHRKFPALFCKAAEAARVLFP----- 186
Db 148 GYDRFVDFHPMFLKVMWGTGSKLYGPTGTDGFRDNQLRFLCLLCLAAEAAPVLNENNS 207
QY 187 ---GP-GEDCVFVANDWHSALVPLVKDEYQPKGFTKRAKSVLAHNIATQGRWEEAFK 242
Db 208 YFSGYGENVVFVANDWHTAVLPCYLKSMYKONGIYVNAKVAFCIHNIAAYOGRFPRVDFE 267
QY 243 DTKLPOAFDKLAFSDGYAKVYTEATPMEDEKPPITGKTYKKNLWKGIIAADKLVTV 302

Db 268 LLNLPSFMSFDFVDGHVK-----PVVG-----RKINWAKGITECDVVLIV 310
QY 303 SPNATEIADAAGVELDVIIRAKGIE-GIVNGMIEEWNPKDKFSLAPYDONSYYAG 361
Db 311 SPHYVELTSGPKGVELDGLVRAKPLETGIVNGMDVVDWNPATDKYISVKNATTVAEA 370
QY 362 KAAAEALQAEGLPVDPTAFATFGRLEEQKGVDIILAAALPKILATPKVQIAILGTGK 421
Db 371 RALNKEILOAEVGLPVDSSIPVIFIGRLEEQKGVDIILAAALPEFL-EENVQIIVLGTGK 429
QY 422 AAYEKLUNAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVPSRFPFPGGLIOLHAMHYGT 481
Db 430 KMEELMLLEAKYQNGARCIKFNPNVLAHMMFAGANFIIVPSRFPFPGGLIOLQGRYGV 489
QY 482 VPVASTGGLVDRVKGSGVGFHMGALNP--DKLDEADADALATVRRASVFFAGGRYPEM 539
Db 490 IPICSTGGLVDIVSEGVTFHMGSNFVEFETVDPADVAASVNVTRALKQYKTPSFHAM 549
QY 540 VANCISQDLWSKPAOKWEGILLEEVYVYKGG-----GVATAKKEIKVP 582
Db 550 VQNCMAQDLWSKPAKW---EAILGLGVESQPGIEETAPLAKONVATP 599

RESULT 4

US-09-952-677-2
; Sequence 2, Application US/09952677
; Patent No. US20020138876A1
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; Lorz, Horst
; Lutticke, Stephanie
; Walter, Lennart
; Froberg, Claus
; Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; FROM WHEAT WHICH ARE INVOLVED IN STARCH
; SYNTHESIS

NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

ADDRESS: James F. Haley, Jr., c/o Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10020

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/952,677
FILING DATE: 14-Sep-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/196,390
FILING DATE: 19-No. US20020138876A1-1998
APPLICATION NUMBER: DE 196 21 588.9
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: DE 196 36 917.7
FILING DATE: 11-SEP-1996
APPLICATION NUMBER: PCT/EP97/02793
FILING DATE: 28-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Haley, Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: AGREVO-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 671 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-952-677-2

Query Match 24.3%; Score 884.5; DB 10; Length 671;
Best Local Similarity 34.3%; Pred. No. 2.3e-54;
Matches 248; Conservative 98; Mismatches 245; Indels 133; Gaps 24;

QY 23 GVKKTANQLRELARGSARKSTSRSAVTGATGATCALD-----IVMVAEAVAP 70
Db 9 GVSELAPDLLE---GIAEDSIDSIIVAASEQDSEIMDANEQPOAKVTRISFVFTGEAP 65
QY 71 WSKTGLGDDVTGGLPIELVKRGHRVMTIAPRY-----DOYADAWDTSVVVDIM---GE 120
Db 66 YAKSGGLGDDVCGSLPITALAARGHVVVMPRYLNGSSDKNYAKALYAKHKKIPCTGCGSH 125
QY 121 KVRYPHSIKKGVHRVHIDHPWFLAKVWGKTSGLYGRSGADYLDNHRKFALEKAAIEA 180
Db 126 EVTFEHRDNDVDFVYDHP-----SYHRPGS-LYGDNFGA-FGDNQFRYTLTLCYAAACEA 178
QY 181 ARVLPFGP---GEDCVFVANDWHSALVPLVKDEYQPKQFTKAKSVLAHNTAFQGRMW 237
Db 179 PLILELGGYIYGQNCMEFVNDWHSALVPLVLAARVYGRVYDRSRSTLVTHNLAHQVGP 238
QY 238 EEAFKDTKLPQAAFDKLAFLSDGYAKVVTATPMEDEKPLTGTGYKKINLKGIIAAD 297
Db 239 ASTYPDGLPPEWVGALW-----VPEWARRHALDXG-----EAVNFKGAVVTAD 285
QY 298 KLVTVSPNATETADAAGVVELDVI--RAKLEGIVNGMDTEENPKTDKFLSAPYDQ 355
Db 286 RIVTVSQGYSEWVTT-AEGQGLNELLSRKSRLNGIVNGIDINDWNPTDKCLPHYSV 344
QY 356 NSVYAKAAAKEALQAEGLPVDPTAPLPAFICRLEEQKGVDIILAAALPKILATPKVQIA 415
Db 345 DDL-SGAKCKAELOKELGLPREDVPLIGTIGRDIYQKIDILIKWAIPELM-REDVQV 402
QY 416 ILGTGKAAVEKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVPSRFPFPGGLIQLH 475
Db 403 MLGSGDPITFEGWNRKSTESSYKDKFRGWGVSVVSHRITAGCDILLMPSEFFECGLNQLY 462
QY 476 AMHYGTVPVASTGGLVDIVK-----EGVTGFHMGALNPKDKLDEADADALATVRR 526
Db 463 AMOYGTVPVYVHTGGLRDIIVTFNPFKAGEGTGWFAPSLTVDKMLWA-----LRT 514
QY 527 ASEVFAGGRYPENVANCISQDLWSKPAOKWEGILLEEVYVYKGVATATKEELKVPVAK 586
Db 515 AMSTFE-----HAPS--WEGLM-----KRGMTKDHITWDH 542
QY 587 IPGDLPAVSYPATNLKPVSAVSEGNAGAAAPKVGTTAPAMGAWRATTPSGPSAAAT---P 643
Db 543 APSSTSRSSSGFSWTNPTSC-----RRGLRSKCESPSALKTSSSSFRGP 587
QY 644 KVTYKPAALPATAKPTAGIK-LAGEASTTSENGAASNGNGNGASAKTSAAKPLVSA 702
Db 588 EGYPTLCRPAIVESQCACLLWFAG-----SRTYDGA-----AAAVTASGGROLQW 635
QY 703 ATRK 706
Db 636 GIRK 639

RESULT 5

US-09-952-677-6
; Sequence 6, Application US/09952677
; Patent No. US20020138876A1
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; Lorz, Horst
; Lutticke, Stephanie
; Walter, Lennart
; Froberg, Claus
; Kossmann, Jens


```

Db 748 GQGFVLGSSFPVPH-----IQREFEGIAHFQNNNIRLLKYYDDSLSHWIAASDMFI 801
QY 462 VPSRFEPCGLIQLHAMHYGVPVYVASTGGGLVDTV-----KEGVTGFHMGALNPKL 512
Db 802 VPSWFPCGLTQMAMRYGSPVVRKTTGGGLNDSYDFDDTTPKELRNGFTF--VHPDE- 858
QY 513 DEADALAAATVRRASEVFAAGRYPEMVANCISODL-----SWSKPAOKWEGILLEEV 565
Db 859 -----KALSGAMERAFNY--NRKPEVWKOLVQKMDKIDFSWASSASQYEDITYORAV 908

RESULT 7
US-10-163-214-2
; Sequence 2, Application US/10163214
; Publication No. US20030097688A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Starch Synthase Isoform V
; FILE REFERENCE: BB1520 US NA
; CURRENT APPLICATION NUMBER: US/10/163,214
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,099
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 909
; TYPE: PRT
; ORGANISM: Zea mays
US-10-163-214-2

Query Match 15.4%; Score 559; DB 9; Length 909;
Best Local Similarity 28.8%; Pred. No. 3.6e-31;
Matches 174; Conservative 98; Mismatches 218; Indels 114; Gaps 23;

QY 10 SSARPVINAASFGVKKTANQLIRELA-----RGSARKSTSRSAVTGA-TG 54
Db 361 SLEKKTINDAS-----MLREMAWKDNKRLREAYLSRRGMEERELIDSPKMLPG 411
QY 55 ATCGLDIVMVAEEVAPWSKGTGGVDTGGLTPELVKRGHRVMTIAPRYD-QYADADWTSV 113
Db 412 TSSGLHIVHIAEAMAPVAKVGGGLADVISGLKALQKKGHLVEIILPKYDCMQHINLNK 471
QY 114 VVDIMGEKRYFHSIKKGVHRVWIDHPWFLAKVWGKT--GSKLY--GPRSGADYL----- 164
Db 472 VLDV-----VVKSYFEGNMFPANKIWTGTVBGLPVYFTEPQHPGKFFWRAQY 517
QY 165 ----DNHKKRALFCKAATEAARVLPFGGEDC-VFVANDWHSALVPVLLKDEYQPKGQFT 219
Db 518 YGEHDDKRFYSFVSVALE----LLYQSGKKVDIICHQDWQAFAPLYWDVYANLG-FN 572
QY 220 KAKSVLAHNTIAFOGRMWEAEAFKDKLPQAFDKLAFSDGYAKVYVTEATPMBEDEKPLT 279
Db 573 SARICFTCHNFYOG-----IAPAQDLAYCGLDVHLDORPDR--MR 611
QY 280 GKTYKKINWLKGGIITAADKLVTSPNYATEIAADAGGVVELDTVIRAKGIEGVINGMDE 339
Db 612 DNSHGRINVGKAVVYSNIVTVPSTYAOEVSEGGRLQDITLVKHSKFFVGLINGIDTD 671
QY 340 ENNPXTDKFLSAPYQNSYAGKAAKALQALGLI-PVDPTAPLFAFAGRLKEEGKGVDI 398
Db 672 TNPSTDFELKVQYSANDLY-CKSANKAALRQLKLASTQASQPLVGCITRLVPOKGVHL 730
QY 399 ILAALPKILATPKVQIATILGTAKAAY-EKLVNAIGTKYK--GRAGVVKFSAPLAHMLTA 455
Db 731 INHAIYKIKTELGG-QVLLGSSPVQHIQREFEGIAHQFQNNNNVRLILLKYDALAHMIFA 789
QY 456 GADEFMLVPSRPEPCGLIQLHAMHYGTVPVVASTGGGLVDT-----VKEGVTGFH 503
Db 790 ASDMFIVPSMEFPCGLTQWAMRYGSPVVRRTGGGLNDSVFDLDDTETIPMEVRNGFTFL- 848

```

```

QY 504 MGALNPKLDRADADALAAATVRRASEVFAAGRYPEMVANCISODL-----SWSKPAOKWEG 559
Db 849 -----KADBDQ---FGNALERAFNY--HRKPEVWKQLVQKMDKIDFSWDTSVSQYEE 896
QY 560 LLEE 563
Db 897 IYQK 900

```

RESULT 8

```

US-10-163-214-12
; Sequence 12, Application US/10163214
; Publication No. US20030097688A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Starch Synthase Isoform V
; FILE REFERENCE: BB1520 US NA
; CURRENT APPLICATION NUMBER: US/10/163,214
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,099
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-163-214-12

```

```

Query Match 15.3%; Score 555.5; DB 9; Length 914;
Best Local Similarity 29.8%; Pred. No. 6.4e-31;
Matches 176; Conservative 96; Mismatches 216; Indels 103; Gaps 24;

QY 31 LLELA-----RGSARKSTSRSAVTGA-TGATCALDIVMVAEEVAPWSKGTG 75
Db 378 MLREMYQKRDSRLREAYLSYRGTEVMDNLLKMLALPGTSSGLHTAHIAEAMAPYAKVG 437
QY 76 GLGDVTTGGPLIELVKRGHRVMTIAPRYD-QYADADWTSVVDIMGEKRYFHSIKKGVHR 134
Db 438 GLADVTSGGLKALQKKGHLVEIILPKYDCMQDVSNLKVLDVLVQS--YF----- 486
QY 135 VWIDHPWFLAKVWGKT--GSKLY--GPR-----SGADYLDNH--KRALFCKAALAEA 181
Db 487 ---EGNMFNKINTGIVEGLPVYFTEPQHPAMFFSRAQYGEHDDKRFYSFVSVALE-- 541
QY 182 RVLPFGGEDC-VFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLAHNTIAFOGRMWEA 240
Db 542 --LLYQSGKKVDIICHQDWQAFAPLYWDVYANLG-FNSARICFTCHNFYOG----- 592
QY 241 FKDTKLPOAFAFKLAFSDGYAKVYVTEATPMBEDEKPP-LTGKTYKKINWLKGGIITAADKL 299
Db 593 -----TAPARDLAWC-----GLDVEHLDRDRMRDNRSHGRINAVKGAVVYSNIV 636
QY 300 VTVPSPNYATEIAADAGGVVELDTVIRAKGIEGVINGMDEEWNPKTDKFLSAPYQNSY 359
Db 637 TTVSPTYALEVRSEGGRLQDITLVKHSKFFVGLINGIDTDWNPCTDRYLKVOYNKADL- 695
QY 360 AKSAAKALQALGLI-PVDPTAPLFAFAGRLKEEGKGVDIILAALPKILATPKVQIATILG 418
Db 696 OGKAAKALRQLKLASTQASQPLVGCITRLVQKGVHLIRHAIYK-TAELGGQFVLG 754
QY 419 TCKAAVEKLVNAIGTKYKGRA-----KGVVFSAPLAHMLTAGADFMVPSRFEPC 469
Db 755 SSP-----VPEIQREFEGIAHFQNNNIRLLKYDDALSCHCIYAAASDMFIVPSIFEPC 808
QY 470 GLIQLHAMHYGTVPVVASTGGGLVDTV-----KEGVTGFHMGALNPKLDRADADALAAATVR 525
Db 809 GLTQMAMRYGSPVVRKTTGGGLNDSVFDLDDTETIP---MEVRNGFTFVKADEQGLSSAME 865

```

QY 526 RASENTPAGGRYPDEWANCISODL-----SWSKPAKWEGLLLEEVVYGKGYA 572
 || : | ||| : : || : || : || : || : || : ||
 Db 866 RAFNCYT--RKPEVKOLQOKMTIDFSWDTSASQEDYYQKAVARAVA 914

RESULT 9
 US-10-163-214-13
 ; Sequence 13, Application US/10163214
 ; Publication No. US20030097688A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Stephen M.
 ; APPLICANT: Broglie, Karen E.
 ; APPLICANT: Butler, Karlene H.
 ; APPLICANT: Thorpe, Catherine J.
 ; TITLE OF INVENTION: Starch Synthese Isoform V
 ; FILE REFERENCE: BB1520 US NA
 ; CURRENT APPLICATION NUMBER: US/10/163,214
 ; CURRENT FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: 60/297,099
 ; PRIOR FILING DATE: 2001-06-08
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 13
 ; LENGTH: 874
 ; TYPE: PRF
 ; ORGANISM: Vigna unguiculata
 US-10-163-214-13

Query Match 15.0%; Score 547.5; DB 9; Length 874;
 Best Local Similarity 29.9%; Pred. No. 2.2e-30;
 Matches 172; Conservative 90; Mismatches 200; Indels 113; Gaps 26;

QY	41	RKSTSRSAVTGATCATCA----	LDIYMVAEAPWSKTGGIGDVTGGLPIELVKRHRV	95
Db	354	KETHEHAISAFGLTSSATSPGLYVHIAEMAPVAKVGSDVISLSKALQKKHLV	413	
QY	96	MTAIPRYD--QYADAWDTSVVVDIMGEKRVYFHSTIKGVHRVHDHPWFIAKVGMT--G	151	
Db	414	EILPKYDCMQYDRIGDLR--ALDVVIES--VF-----DGQLFKNKIWGTVEG	458	
QY	152	SKLY-----	GPRSGADYLNHKRFAFLCFAAIPTAARVLPTGPGEDCVFVAND	198
Db	459	LPVYFIIEPHPGKFFWRGDYYGAH--DDFRFSYFSRAALEF--ILOAGKQPDIH--CHD	513	
QY	199	WHSALVPULLKDEYOPKGOFTKAKSLVLAHNIAFPQRMWEAFKDTKLPOAFAFDKLAESD	258	
Db	514	WQTAFIAPLYWDVYAPKG--LNSARICFTCHNFEOGTAGASELEACGLDSHOLNR----	567	
QY	259	GYAKVYTEATPMEDEKPPLGKYKKLNWLKGLIIAADKLVTSPNTATEIADAAGV	318	
Db	568	-----PDRMQDN-----SAHNRVNSYKGVAVVYNIWTVTSPTYAQEVRT--AEGGK	611	
QY	319	ELDTVI--RANGIEGIIVNGMDEENPNKTDKFLSAPYDONSYAGKAAKEALOAEGL-	375	
Db	612	GLHSTLSHKXKFILNGITDDINWPATDFLOQVYANDL-QQSENKEALRNGLS	670	
QY	376	PVDPTAPLFAFIGRLEEQKGVDIILAALPKILATPKV--QIAILGTGAAAYEKLVNAIGT	433	
Db	671	SADVRRPLVGCITRLVPQKGLIRHA---IYLTLELGQQFVLLGSSVPVH-----IQR	721	
QY	434	KYKGRA-----	KGVVKSAPTAAHMLTAGADFMLVPSRPREGCLITOLHAMHYGTVPV	484
Db	722	EFEGIANHFQNHDHTRLILKYDESLSHAIYAASDMFIIPSIFEPGCLTQMISMRYGAPI	781	
QY	485	VASTGGLVDV-----	KEGVTFGHMGALNPDKLDDADADAATAATVRRASEVFA	532
Db	782	ARKUGLNDSEVDVDDDTIPSFQRNGT-----FLNAD--EKGINDALV----RAINLFT	830	
QY	533	GG--RYPEMVANCISODLSWSKPAKWEGLLLEEV	565	
Db	831	NDPKSNKOLVQKMDMNIDFSWDSAAQYEELYSKVS	865	

RESULT 10

US-10-007-693-99

Sequence 99, Application US/10007693

Patent No. US20020146776A1

GENERAL INFORMATION:

APPLICANT: Bhatia, Ajay

APPLICANT: Probst, Peter

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT

TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION

FILE REFERENCE: 210121.515C2

CURRENT APPLICATION NUMBER: US/10/007,693

CURRENT FILING DATE: 2001-12-05

NUMBER OF SEQ ID NOS: 157

SEQ ID NO 99

LENGTH: 474

TYPE: PRT

ORGANISM: Chlamydia trachomatis serovar D

US-10-007-693-99

Query Match 11.3%; Score 410.5; DB 12; Length 474;

Best Local Similarity 28.2%; Pred. No. 4.5e-21;

Matches 148; Conservative 72; Mismatches 206; Indels 99; Gaps 22;

QY 59 LDIWVAAEVPWKSGTGLGSDVTGGLPTELKGRHVMTIAPRDQVADAWDTSVVDIM 118

DB 1 MKIIITALEEFAPIKAGGLDALYGL-AKALAAHNTTEVILPKFTLTPKEQDLCSI- 58

QY 119 GEKYRYF-----HSIKKGVRHVWIDHPWFLAKVWGKTSKLYGPRSGADYL 164

DB 59 -OKUSYFPAGQENATFSYFEGIKVTLFKL-----DTQPELENAETIYS 104

QY 165 DNHKRFALFCKAAI-----EAARVLPGFGECEVFNANDHSALVPVLLKDEYQKQFT 219

DB 105 DDAFFECASAAAASYIQKEGANIVHL-----HDWHTGLVAGLLKQ--QPCSQLQ 152

QY 220 KAKSVLAIHNTAFQGRMEEAFKOTKLPOAAFDKLAFLSDGYAKVYVTEATPMEDEKPLT 279

DB 153 KI-VLTUHNFGYRGYTTREILEASSLNE-----FYISQYQIFRDPQTCVL 196

QY 280 KTKYKINLWKGIIIAADKLVTVSPNATEIAADAAGVEIDTVIRAK--GIEGIVNGMD 337

DB 197 -----LKGALYCSDFVTVTSYAKILEEDYS-DYEIHDALTARQHHLRGILNGID 246

QY 338 IEENPKTKLSADYDQ-----NSVYAGRAAKAEALQAEGLPVDPTAPLFAFIGLE 391

DB 247 TTWGPETDPNLAKNYTKELFETPSIFFEAKAENKALYERIGLSLE-HSPCVCIIISRIA 305

QY 392 EKGQGDITLALPKILATPKVQIATLGH--GKAAVEKLVN---AIGTKYKGRAGKAWKFS 446

DB 306 EKGQGFHKQAILHLEAENAYILI-IIGCYGNQLHEEFANQLQESLANSFDVRI--LLIYS 362

QY 447 APLAHLMTAGADFMVLVSRPFECGLIQLHAMHYGVTVVAVSTGGVLVDIVKEGVTFHMA 506

DB 363 DVLARQIFAADMCIPSMFECGLTQMIGMRYGVPLVRATFGGLADVANGINGFSF-- 420

QY 507 LNP-DKLEADADALAAVRRASEVFAGRYPEMVANCI--SODL 548

DB 421 FNPFDYEFERNMLSEAVTYRTNH-----DKWHQIIVRACLDFSSDL 461

RESULT 11

US-10-007-693-69

Sequence 69, Application US/10007693

Patent No. US20020146776A1

GENERAL INFORMATION:

APPLICANT: Bhatia, Ajay

APPLICANT: Probst, Peter

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT

TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION

FILE REFERENCE: 210121.515C2

CURRENT APPLICATION NUMBER: US/10/007,693

CURRENT FILING DATE: 2001-12-05

NUMBER OF SEQ ID NOS: 157

```
; SEQ ID NO 69
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-007-693-69

Query Match      11.1%; Score 402.5; DB 12; Length 476;
Best Local Similarity 26.9%; Pred. No. 1.7e-20;
Matches 150; Conservative 78; Mismatches 204; Indels 125; Gaps 24;

QY 61 IVWVAEAPMSKTTGGGLDVTGGPIELVKRHRVMTIAPRY---DOVADAWDTSVVVDI 117
Db 3 IVQVAVEFTPIVKVGGGLDVAASLSKELAKQ-NDVEVLLPHYPLISFESS-----QV 54
QY 118 MGEKVRVFHSKKGVHRVWDHPWFLAKVWCKTKGSKLYGP-----RSGAD 162
Db 55 LSENSFYE-----FLGKQOASALSISYEGTLTIITLDSQIELFSTTSV 99
QY 163 YLDNH-KRFALFCKAAIEAARVLPFGPGEDCFVANDWHSALVPVLKDKXQPKQFTKA 221
Db 100 YSENNVRFSAFAAAA--AAYLQADPAD--IVHLHDWHVGLLAGLLKNPLNP-----VHS 151
QY 222 KSVLAHINIAFQGRWEEAEKDTKLQAAFDKLAFSGDYAKVYTEATPMEDEKPLPTGK 281
Db 152 KIVFTIHFYRG-----YCSQTLLAAS-----QIDDFHLSHY-----QLFRDPQTSV 194
QY 282 TYKKNWLGKGIITAAADKLVTSPNYATEIAADAAGGVELDTVIRAKG--LEGIVNGMDIE 339
Db 195 L-----MKGALYCSYITVSLTYQEIINDYS-DYELHDAILARNSVFSGIINGDED 247
QY 340 ENPKTKDFLSAPYD-----QNSVYAGKAAAEALQAEGLPVDPTAPLFAFGRLEEQ 393
Db 248 VWNPKTDPALAVQYDASLLSEPDVLTKEENRAVLYEKLGISSD-YFPLICVLSRIVEE 306
QY 394 KG-----VDIILALPKILATPKVOIALGTGKAAEKLVAIGTKYKGRAG-----VVK 444
Db 307 KGPEFKEIILHMEHSYA-----FILIGTSQ--NEVLNFEFRNIQDCLASSPNRILLD 359
QY 445 FSAPLAHMTAGADEMLVPSREFPCGLLIQHLHAMHYGTVPVASTGGGLVDTVKGVGFHM 504
Db 360 FNDPLARLYAADMCIPIFSHREACGLTQILAMRYGTVPVLRKVTGGGLADTVIPGVNGFTF 419
QY 505 GALNPDKLDEADALAAATVRRASEVFAAGRYPEMVANCISODLSNKSPPAKWEGGLEEV 564
Db 420 --FDTNNE-----FRAMLSNAVY-----TYRQEPDVLNLIESG 453
QY 565 VYKGGVATAKKEIKV 581
Db 454 MLRASGLDAMAKHYVNL 470

RESULT 12
US-10-163-214-10
; Sequence 10, Application US/10163214
; Publication No. US20030097688A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Starch Synthase Isoform V
; FILE REFERENCE: BB1520 US NA
; CURRENT APPLICATION NUMBER: US/10/163,214
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,099
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-163-214-10

Query Match      11.1%; Score 402.5; DB 12; Length 476;
Best Local Similarity 26.9%; Pred. No. 1.7e-20;
Matches 150; Conservative 78; Mismatches 204; Indels 125; Gaps 24;

QY 61 IVWVAEAPMSKTTGGGLDVTGGPIELVKRHRVMTIAPRY---DOVADAWDTSVVVDI 117
Db 3 IVQVAVEFTPIVKVGGGLDVAASLSKELAKQ-NDVEVLLPHYPLISFESS-----QV 54
QY 118 MGEKVRVFHSKKGVHRVWDHPWFLAKVWCKTKGSKLYGP-----RSGAD 162
Db 55 LSENSFYE-----FLGKQOASALSISYEGTLTIITLDSQIELFSTTSV 99
QY 163 YLDNH-KRFALFCKAAIEAARVLPFGPGEDCFVANDWHSALVPVLKDKXQPKQFTKA 221
Db 100 YSENNVRFSAFAAAA--AAYLQADPAD--IVHLHDWHVGLLAGLLKNPLNP-----VHS 151
QY 222 KSVLAHINIAFQGRWEEAEKDTKLQAAFDKLAFSGDYAKVYTEATPMEDEKPLPTGK 281
Db 152 KIVFTIHFYRG-----YCSQTLLAAS-----QIDDFHLSHY-----QLFRDPQTSV 194
QY 282 TYKKNWLGKGIITAAADKLVTSPNYATEIAADAAGGVELDTVIRAKG--LEGIVNGMDIE 339
Db 195 L-----MKGALYCSYITVSLTYQEIINDYS-DYELHDAILARNSVFSGIINGDED 247
QY 340 ENPKTKDFLSAPYD-----QNSVYAGKAAAEALQAEGLPVDPTAPLFAFGRLEEQ 393
Db 248 VWNPKTDPALAVQYDASLLSEPDVLTKEENRAVLYEKLGISSD-YFPLICVLSRIVEE 306
QY 394 KG-----VDIILALPKILATPKVOIALGTGKAAEKLVAIGTKYKGRAG-----VVK 444
Db 307 KGPEFKEIILHMEHSYA-----FILIGTSQ--NEVLNFEFRNIQDCLASSPNRILLD 359
QY 445 FSAPLAHMTAGADEMLVPSREFPCGLLIQHLHAMHYGTVPVASTGGGLVDTVKGVGFHM 504
Db 360 FNDPLARLYAADMCIPIFSHREACGLTQILAMRYGTVPVLRKVTGGGLADTVIPGVNGFTF 419
QY 505 GALNPDKLDEADALAAATVRRASEVFAAGRYPEMVANCISODLSNKSPPAKWEGGLEEV 564
Db 420 --FDTNNE-----FRAMLSNAVY-----TYRQEPDVLNLIESG 453
QY 565 VYKGGVATAKKEIKV 581
Db 454 MLRASGLDAMAKHYVNL 470

Query Match      5.9%; Score 215.5; DB 10; Length 117;
Best Local Similarity 46.6%; Pred. No. 3.9e-08;
Matches 41; Conservative 17; Mismatches 25; Indels 5; Gaps 1;

QY 413 QIALIGTG-----KAAVEKLVAIGTKYKGRAGVVKFSAPLAHMTAGADEMLVPSREF 467
Db 8 QVVLGSAIPHRIQDFTNLASKLHGEYHGVKLCITYDEPLSHLIYAGADFILVPSMFE 67
QY 468 PCGLIQLHAMHYGTVPVASTGGGLVDTV 495
Db 68 PCGLTQLTAMRYGSIPIVRKVTGGGLYDTV 95

RESULT 14
US-09-738-626-4738
; Sequence 4738, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
```

```

; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4738
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4738

Query Match 5.7%; Score 206; DB 9; Length 409;
Best Local Similarity 27.2%; Pred. No. 1e-06;
Matches 83; Conservative 47; Mismatches 127; Indels 48; Gaps 15;

QY 267 ATPMERDE--KPPLTGCTKYYKINWL-KGGLIAADKLVVSPNVTATEIADAAGVELDTV 323
Db 129 AHSLEPDRPWKRELGGYDVSSSEKNAMEYADAVIANVARMKDSILA-AYPRIEPDNV 187
QY 324 IRAKGTIEGVNGDIEENPKTDKFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPL 383
Db 188 -----RVVLNGIDTLMQPR-----PTFDD-----AEDSVLRSLG--VDQRPPI 224
QY 384 FAFIGRLEEQGVDIILALPKILATPKVQIATLIGTKAAYE--KLVNAIGTKYKGRAG 441
Db 225 VAFVGRITRQKGVHELIKA--AALFDESQVLVLCAGAPDPTETAARTTALVEELQAKREG 282
QY 442 VVRFSAPLA-----HMLTAGADFMVLPSPREPGLQLHMHYGTVPVAVASTGGGLVTVK 496
Db 283 IFWVQDMGLKDKIQEILTA-ADTFVCPSTYIEPLGIVNLEAMACNTAVVASDVGGIPEVVV 341
QY 497 EGVGTG--FHMALNPKLDEADADA---LAATVRRASEVFAGGRYPFWVANCISQDLSWS 551
Db 342 DGTGALVHYDENDVETFERDIAEAVNVKAVDRETAAKFGLAGR--ERAIN-----DFSWA 395
QY 552 KPAQK 556
Db 396 TIAQO 400

RESULT 15
US-09-738-626-5896
; Sequence 5896, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4738
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-5896

Query Match 5.2%; Score 189; DB 9; Length 381;
Best Local Similarity 27.0%; Pred. No. 1.5e-05;
Matches 74; Conservative 43; Mismatches 95; Indels 62; Gaps 14;

QY 317 GVVELD-----TVIRAKG-----IEGIVNGMDIEENPKTDKFLSAPYDQNSVYAG 361
Db 138 GTEVDVLTYSQYTLRRFRKSAFGSHPTFEHLPSGVDVKRTPAT-----PED----- 184
QY 362 KAAAKEALQAEGLPVDPTAPLEAFICRLEEQGVDIILALPKILAT-PKVQIATLIGTC 420
Db 185 -----KSATRKKLGF--TDTPVIACNSRLVPRKQDLSLIKAMPQVIAARPDQALLIVGSG 238
QY 421 KAAAYEKLIVNAIGTKYKGRAGVVKFSAPLAHM-----LTAGADFMVLPSPR-----FEPC 469
Db 239 R--YESTLRLATDVSON-----VKFLGRLEYQDMINTIAAADIFAMPARTRGGLDVEGL 292
QY 470 GLIQLHAMHYGTVPVAVAST--GGLVDVTYKGVTFHMGALNPKLDEADADAATVRRAS 528
Db 293 GIVYLEAQACG--VFVIAAGTSGGAPETVTP--ATGLVVEGSDVKLSELLIELLDPIRRAA 350
QY 529 EVFAGGRYPFWVANCISQDLSWSKPAQKWEGLLE 562
Db 351 MGRAGRAH-----VEAENSWEIMGERLINILQ 377

Search completed: June 4, 2003, 15:15:16
Job time : 27.8651 secs
```

```

; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4738
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4738

Query Match 5.7%; Score 206; DB 9; Length 409;
Best Local Similarity 27.2%; Pred. No. 1e-06;
Matches 83; Conservative 47; Mismatches 127; Indels 48; Gaps 15;

QY 267 ATPMERDE--KPPLTGCTKYYKINWL-KGGLIAADKLVVSPNVTATEIADAAGVELDTV 323
Db 129 AHSLEPDRPWKRELGGYDVSSSEKNAMEYADAVIANVARMKDSILA-AYPRIEPDNV 187
QY 324 IRAKGTIEGVNGDIEENPKTDKFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPL 383
Db 188 -----RVVLNGIDTLMQPR-----PTFDD-----AEDSVLRSLG--VDQRPPI 224
QY 384 FAFIGRLEEQGVDIILALPKILATPKVQIATLIGTKAAYE--KLVNAIGTKYKGRAG 441
Db 225 VAFVGRITRQKGVHELIKA--AALFDESQVLVLCAGAPDPTETAARTTALVEELQAKREG 282
QY 442 VVRFSAPLA-----HMLTAGADFMVLPSPREPGLQLHMHYGTVPVAVASTGGGLVTVK 496
Db 283 IFWVQDMGLKDKIQEILTA-ADTFVCPSTYIEPLGIVNLEAMACNTAVVASDVGGIPEVVV 341
QY 497 EGVGTG--FHMALNPKLDEADADA---LAATVRRASEVFAGGRYPFWVANCISQDLSWS 551
Db 342 DGTGALVHYDENDVETFERDIAEAVNVKAVDRETAAKFGLAGR--ERAIN-----DFSWA 395
QY 552 KPAQK 556
Db 396 TIAQO 400

RESULT 15
US-09-738-626-5896
; Sequence 5896, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
```


GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:49:45 ; Search time 34.5277 Seconds
(without alignments)
2512.365 Million cell updates/sec

Title: US-09-980-771a-3_COPY_58_708

Perfect score: 3370

Sequence: 1 ALDINVAEAPWPKTGL.....SASKTSAKPLYSAAETKSA 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1440	42.7	606	13 AAR25476	Granule-bound star
2	1397.5	41.5	609	19 AAW56485	Oryza sativa starch
3	1390	41.2	527	21 AAG04668	Arabidopsis thalia
4	1390	41.2	563	21 AAG04667	Arabidopsis thalia
5	1390	41.2	610	21 AAG04666	Arabidopsis thalia
6	1390	41.2	610	23 ABB91229	Arabidopsis thalia
7	1353.5	40.2	533	19 AAW56484	Herbicidally activ
8	1344.5	39.9	637	19 AAY16604	Zea mays waxy gene
9	1331.5	39.5	502	13 AAR25474	Protein encoded by
10	1194	35.4	238	22 AAB31175	Central fragment f

11	904	26.8	534	21 AAB49307	Wheat starch synth
12	898	26.6	792	23 ABB92160	Herbicidally activ
13	891.5	26.5	647	20 AAY09004	Wheat starch solub
14	882.5	26.2	649	19 AAW38218	Maize starch synth
15	881	26.1	671	19 AAW39337	Wheat soluble star
16	881	26.1	756	21 AAY50818	Wheat soluble star
17	873.5	25.9	626	15 AAR51231	Soluble rice starch
18	870.5	25.8	652	23 ABB93595	Herbicidally activ
19	861.5	25.6	539	19 AAW56491	Zea mays starch sy
20	861.5	25.6	583	19 AAW56488	Zea mays soluble s
21	852.5	25.3	583	19 AAW70894	Maize starch solub
22	847	25.1	802	23 AAU99845	Modified barley st
23	847	25.1	813	23 AAU99844	Barley cultivar Mo
24	846	25.1	812	23 AAU99847	Barley line MK6827
25	845	25.1	804	19 AAW70892	Maize starch solub
26	844	25.0	798	21 AAB37566	Wheat starch synth
27	842	25.0	799	21 AAB37567	Wheat starch synth
28	842	25.0	799	21 AAB37597	Wheat starch synth
29	841	25.0	799	21 AAW23938	Wheat granule-boun
30	837	24.8	698	19 AAW56487	Zea mays soluble s
31	825.5	24.5	812	23 AAU99846	Barley line 292 st
32	825	24.5	597	21 AAB37568	Wheat starch synth
33	767	22.8	466	23 AAR51855	Rice starch synth
34	726	21.5	459	17 AAR99540	Soluble starch syn
35	701.5	20.8	669	19 AAW70893	Maize starch solub
36	701.5	20.8	669	19 AAW56486	Zea mays soluble s
37	669	19.9	483	23 ABB54012	Lactococcus lactis
38	647.5	19.2	476	23 ABP27592	Streptococcus poly
39	580	17.2	477	13 AAR25462	giga. Escherichia
40	580	17.2	477	19 AAW53890	E. coli glycochen s
41	568.5	16.9	677	17 AAR99539	Soluble starch syn
42	568.5	16.9	1230	18 AAW17785	Potato tuber solub
43	568.5	16.9	1230	21 AAB49306	Potato starch synt
44	561.5	16.7	1025	23 ABB90967	Herbicidally activ
45	546	16.2	495	19 AAW70885	Protein encoded by

ALIGNMENTS

RESULT 1
AAR25476
ID AAR25476 standard; Protein; 606 AA.
XX
AC AAR25476;
XX
DT 15-JAN-1993 (first entry)
XX
DE Granule-bound starch synthase of potato.
XX
KW GBSS gene; amylopectin; amylose production;
KW inhibition; promoter; antisense construct.
XX
OS Solanum tuberosum.
XX
PN WO9211376-A.
XX
PD 09-JUL-1992.
XX
PF 20-DEC-1991; 91WO-SE00892.
XX
PR 21-DEC-1990; 90SE-0004096.
XX
PA (AMYL-) AMYLOGENE HB.
XX
PI Hofvander P, Persson PT, Tallberg A, Wikstrom O;
XX
DR WPI, 1992-250096/30.
XX
DR N-PSDB; AAR26404.
XX
PT Modifying potatoes to form amylopectin starch - using an
PT anti-sense construct to inhibit granule-bound starch synthase

PS Claim 6; Page 28-33; 46pp; English.

XX A genomic library in EMBL3 was prepared using leaves of the potato.

CC Bintje. The library was screened with cDNA clones for the 5' and 3'

CC ends of the GBSS gene. A full-length clone of potato GBSS gene was

CC identified (wx311) and isolated from the genomic library. The gene

CC contained 12 introns. The amino acid sequence was deduced from it.

CC Three fragments of the full-length sequence (5'-end, middle and

CC 3'-end sequences) were characterised as suitable for use in novel

CC antisense constructs to suppress amylose formation in potatoes.

CC See also Q26400-3.

XX SQ Sequence 606 AA;

Query Match 42.7%; Score 1440; DB 13; Length 606;

Best Local Similarity 53.5%; Pred. No. 9.2e-105;

Matches 292; Conservative 67; Mismatches 145; Indels 42; Gaps 10;

QY 2 LDIVVAAAEVAPWSTGGLGDTGGLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDI- 60

DB 81 MNLFVTEGVPSWKTGELGDLGLPPALARGHRVMTISPRYDQYKADWDTGVAVEVK 140

QY 61 MG---EKVYRHSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHHKFALEFC 117

DB 141 VGDSEIIVRFFHCYKRGVDRVDFHPFLEKVMGKTSKLYGPRSGADYLDNHHKFALEFC 200

QY 118 KAAIEAARVLP-----GP-GEDEVFVANDHSAVLPVLLKDEYQPKQFTKAKSVLA 169

DB 201 QALAEAPKVLNLSNYSFGPGEDVLFANDWHALIPCYLKSVMYQSGIYLNAKVAFC 260

QY 170 IHNIAFGRMWEAEAFKDTLPQAAFDKLAFCSDGYAKVYTPATPMEDEKPPITGKTYKKI 229

DB 261 IHNIAFGRMWEAEAFKDTLPQAAFDKLAFCSDGYAKVYTPATPMEDEKPPITGKTYKKI 302

QY 230 NMLKGLIADKLVTPSPNVAETAAAGVVELDTVIRAKGIEGIVNGMDIEEWNPKTD 289

DB 303 NMLKGLIADKLVTPSPNVAETAAAGVVELDTVIRAKGIEGIVNGMDIEEWNPKTD 362

QY 290 KELSAPYDQNSVYAGKAAKAEALQAEGLPVPDPTAPLFAFIRGLEPKQGVDIILAAALPKI 349

DB 363 KYTDVKYDITVMDAKPLKEALQAAVGLPVDKIPILGIFIGLEPKQGVDIILAAALPKI 422

QY 350 LATPKYQIATLGTKAAAYEKLVAIGTKYGRAGVVKFSAPLAHMLTAGADPMLVPSRF 409

DB 423 IGL-DVQIVVLTGKTFEKEFEQIEQLVLYPNKAKGVAKFNVPPLAHMLTAGADPMLVPSRF 481

QY 410 EPCGLLIQHAMHYGTVPVASTGGLVDTKVEGTVGFHMGALNP--DKLDEADADALAATV 467

DB 482 EPCGLLIQHAMHYGTVPVASTGGLVDTKVEGTVGFHMGALNP--DKLDEADADALAATV 541

QY 468 RRASEVFAGGRYPPEMVANCISQDLSWSKPAQKWEGLL-----EEVYKGGVATAKK 519

DB 542 ARALAVYGLTAFEMIKNCMSBELSKPAKKEWETLLGLGASGSEPGVGE-EIAPLAK 600

QY 520 EETKVP 525

DB 601 ENVATP 606

RESULT 2

AAW56485

ID AAW56485 standard; Protein: 609 AA.

XX AAW56485;

AC AAW56485;

XX AAW56485;

DI 11-SEP-1998 (first entry)

XX Oryza sativa starch (bacterial glycogen) synthase.

DE SER: starch-encapsulating region; fusion vector;

XX starch synthase; bacterial glycogen.

KW starch synthase; bacterial glycogen.

XX Oryza sativa.

OS

XX WO9814601-A1.

XX 09-APR-1998.

XX 30-SEP-1997; 97WO-US17555.

XX 30-SEP-1996; 96US-0026855.

XX (EXSE-) EXSEED GENETICS LLC.

XX Guan H, Keeling P;

XX WPI; 1998-240100/21.

XX N-PSDB; AAV29753.

XX Hybrid polypeptide comprising starch-encapsulating region and

PT protein - useful for, e.g. producing protein(s) resistant to

PT degradation by stomach acids

XX

XX Example 2; Page 32; 156pp; English.

XX The sequence is that of starch (bacterial glycogen) which is

CC encoded by the waxy gene. It can be used in the production of a hybrid

CC polypeptide comprising a starch-encapsulating region (SER) fused

CC to a payload protein. The hybrid polypeptide can be used to make

CC modified starches comprising the payload protein, selected from,

CC e.g. hormones, growth factors, antibodies, enzymes, dyes,

CC immunoglobulins, etc. The modified starch can also be used

CC to provide grain feeds enriched in amino acids. By encapsulating

CC the payload protein in starch, it is more resistant to

CC degradation by stomach acids.

XX SQ Sequence 609 AA;

Query Match 41.5%; Score 1397.5; DB 19; Length 609;

Best Local Similarity 51.4%; Pred. No. 2.1e-101;

Matches 281; Conservative 76; Mismatches 147; Indels 43; Gaps 9;

QY 2 LDIVVAAAEVAPWSTGGLGDTGGLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDI- 60

DB 83 MNVFEVGAEMAPWSTGGLGDLGLPPALARGHRVMTISPRYDQYKADWDTSVVAEIK 142

QY 61 ---MGEKVRYSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHHKFALEFC 117

DB 143 VADRYEKVRFFHCYKRGVDRVDFHPFLEKVMGKTSKLYGPRSGADYLDNHHKFALEFC 202

QY 118 KAAIEAARVLP-----PFGP-GEDEVFVANDHSAVLPVLLKDEYQPKQFTKAKSVLA 169

DB 203 QAALEAPRILNLSNYSFGPGEDVLFANDWHALIPCYLKSVMYQSGIYLNAKVAFC 262

QY 170 IHNIAFGRMWEAEAFKDTLPQAAFDKLAFCSDGYAKVYTPATPMEDEKPPITGKTYKKI 229

DB 263 IHNISYQGFAPEDYPEINLSERFSRFSDFIDGY-----DTPVEG-----RKI 305

QY 230 NMLKGLIADKLVTPSPNVAETAAAGVVELDTVIRAKGIEGIVNGMDIEEWNPKTD 289

DB 306 NMLKGLIADKLVTPSPNVAETAAAGVVELDTVIRAKGIEGIVNGMDIEEWNPKTD 365

QY 290 KELSAPYDQNSVYAGKAAKAEALQAEGLPVPDPTAPLFAFIRGLEPKQGVDIILAAALPKI 349

DB 366 KYITAKYDATTAEKALKEALQAEAGLPVDRKIPLIAFIRGLEPKQGVDIILAAALPKI 425

QY 350 LATPKYQIATLGTKAAAYEKLVAIGTKYGRAGVVKFSAPLAHMLTAGADPMLVPSRF 409

DB 426 M-QEDVQIVLTGKTFEKEFEQIEQLVLYPNKAKGVAKFNVPPLAHMLTAGADPMLVPSRF 484

QY 410 EPCGLLIQHAMHYGTVPVASTGGLVDTKVEGTVGFHMGALNP--DKLDEADADALAATV 467

DB 485 EPCGLLIQHAMHYGTVPVASTGGLVDTKVEGTVGFHMGALNP--DKLDEADADALAATV 544

QY 468 RRASEVFAGGRYPPEMVANCISQDLSWSKPAQKWEGLL-----VATAK 518

Db 545 KRAIKVGTAYEMVRNMCNODLSWKGPAKNWNL--LGLGVAGSAPGEGDEIAPLA 602
QY 519 KEEIKVP 525
Db 603 KENVAAP 609

RESULT 3
AAG04668
ID AAG04668 standard; Protein; 527 AA.
XX AC AAG04668;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 782.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-01334256.
PR 14-MAY-1999; 99US-013418.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.


```
PR 22-JUN-1999; 99US-0139899;
PR 23-JUN-1999; 99US-0140353;
PR 23-JUN-1999; 99US-0140354;
PR 24-JUN-1999; 99US-0140695;
PR 28-JUN-1999; 99US-0140823;
PR 29-JUN-1999; 99US-0140991;
PR 30-JUN-1999; 99US-0141287;
PR 01-JUL-1999; 99US-0141842;
PR 01-JUL-1999; 99US-0142154;
PR 02-JUL-1999; 99US-0142055;
PR 06-JUL-1999; 99US-0142390;
PR 08-JUL-1999; 99US-0142803;
PR 09-JUL-1999; 99US-0142920;
PR 12-JUL-1999; 99US-0142977;
PR 13-JUL-1999; 99US-0143542;
PR 14-JUL-1999; 99US-0143624;
PR 15-JUL-1999; 99US-0144005;
PR 16-JUL-1999; 99US-0144085;
PR 16-JUL-1999; 99US-0144086;
PR 19-JUL-1999; 99US-0144325;
PR 19-JUL-1999; 99US-0144331;
PR 19-JUL-1999; 99US-0144332;
PR 19-JUL-1999; 99US-0144333;
PR 19-JUL-1999; 99US-0144334;
PR 19-JUL-1999; 99US-0144335;
PR 20-JUL-1999; 99US-0144352;
PR 20-JUL-1999; 99US-0144632;
PR 20-JUL-1999; 99US-0144884;
PR 21-JUL-1999; 99US-0144814;
PR 21-JUL-1999; 99US-0145086;
PR 21-JUL-1999; 99US-0145088;
PR 22-JUL-1999; 99US-0145085;
PR 22-JUL-1999; 99US-0145087;
PR 22-JUL-1999; 99US-0145089;
PR 22-JUL-1999; 99US-0145192;
PR 23-JUL-1999; 99US-0145218;
PR 23-JUL-1999; 99US-0145218;
PR 23-JUL-1999; 99US-0145224;
PR 26-JUL-1999; 99US-0145276;
PR 27-JUL-1999; 99US-0145913;
PR 27-JUL-1999; 99US-0145918;
PR 27-JUL-1999; 99US-0145919;
PR 28-JUL-1999; 99US-0145951;
PR 02-AUG-1999; 99US-0146386;
PR 02-AUG-1999; 99US-0146388;
PR 03-AUG-1999; 99US-0147038;
PR 04-AUG-1999; 99US-0147204;
PR 04-AUG-1999; 99US-0147302;
PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149929;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;

PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 14-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 22-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161992;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

Query Match 41.2%; Score 1390; DB 21; Length 610;
Best Local Similarity 52.7%; Pred. No. 8.2e-101;
Matches 287; Conservative 72; Mismatches 146; Indels 40; Gaps 12;

QY 2 LDIVAAEYAPWSTGGGLDVTGGPLIELVGRGHRVMTIAPRYQYADMTSVVVDI- 60
Db : : : ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84 MSVIFGAEGVGPWSTGGGLDVLGGLPPALAAARGHRVMTICPRDYQDKAWDTCVVVQK 143
QY 61 MGEK---VRYFHSIKKGHRVWIDHPWFLAKYVGKTKGLYGRSGADYLDNKKRFLFC 117
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
144 VGDKVENVRFFCYKRGVDRVFDHPFLAKYVGKTKGLYGPITGVDDYNDNQLRFLSLC 203
QY 118 KAAIEAARVLPF-----GP-GEDEVFVANDWHSALVPLVLKDEYQPKQFTKAKSVLA 169
Db :|||:||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
204 QAALEAPQVILNLSKSYFGPYGEDVVFVANDWHTALLPCYLKSMYQSGVYNNKAVVFC 263
QY 170 IHNIAFOGRMEEAFKDTKLP---QAAFDKLAESDGYAKVYTEATPMEDEKPLTGKTY 226
Db |||||:||| : ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| |||
264 IHNIAOGRFAFDYSLNLSKSYFGPYGEDVVFVANDWHTALLPCYLKSMYQSGVYNNKAVVFC 303
QY 227 KKNWLKGGIITAADKLVTVSPNYATEIAADRAAGGVELDTVIRAKIEGIVNGMDTEENP 286
Db :|||:||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
304 RKNMMAAILEAHEVLTSPNYAAGELISGVDRGVDELKYLKMKTVSGIINGMDVQEWNP 363
```

QY 287 KTDKFLSAPYDQNSYAGKAAKEALQAEGLGVPDPTAFPLFAFAGRLSEQKGVDIILAL 346
 Db 364 STDKYIDIKYDITVTDAKPLIKEALQAAVGLPVDVDPVVICFICGLEEQKGSILVEAI 423
 QY 347 PKLATPKVQJAILGTGKAAEKLVAIGTKYKGRAGVGVKFSAPLAHMLTAGADPMLVP 406
 Db 424 SKFMGL-NVQMVLGTGKAAEKLVAIGTKYKGRAGVGVKFSAPLAHMLTAGADPMLVP 482
 QY 407 SRFEPGGLIQLHAMHYGTVPVASTGGLVDTVKEGVTGFHMGALNP--DKLDEADADALA 464
 Db 483 SRFEPGGLIQLHAMHYGTVPVASTGGLVDTVKEGVTGFHMGALNP--DKLDEADADALA 542
 QY 465 ATVRASEVFAGGRYPYEMVANCISQDLSWSKPAQKWEGLLEEVYKGVATAKKEIKV 524
 Db 543 KAVTRAVAVYGTSAQEMVKNKMDQDFSNKGPAPLWELKVLNVA-GSEAGTEGEEI-A 600
 QY 525 PVAEK 529
 Db 601 PLAKE 605

RESULT 6
 ABB91229 standard; Protein; 610 AA.
 AC ABB91229;
 DT 31-MAY-2002 (first entry)
 XX Herbicidally active polypeptide SEQ ID NO 440.
 DE Herbicidal; plant; agriculture; herbicide.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 XX WO200210210-A2.
 PN 07-FEB-2002.
 XX 28-AUG-2001; 2001WO-EP09892.
 XX 28-AUG-2001; 2001WO-EP09892.
 PR (FARB) BAYER AG.
 PA Tietjen K, Weidner M;
 PI WPI; 2002-269010/31.
 DR Identifying plant target proteins for herbicidally active compounds,
 XX comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX Claim 5; SEQ ID NO 440; 26lpp + Sequence Listing; English.
 PS The invention relates to identifying target proteins
 XX (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 XX Sequence 610 AA;
 SQ Query Match 41.2%; Score 1390; DB 23; Length 610;
 Best Local Similarity 52.7%; Pred. No. 8.2e-101;
 Matches 287; Conservative 72; Mismatches 146; Indels 40; Gaps 12;

QY 2 LDIWVAEAVPWSKGTGGLGDTGGLPIELVKGHRVMTIAPYDQYADAWDTSVVVDI-60

Db 84 MSVIFIGAEVGPWSKGTGGLGDLVGLPPLAARGHRVMTICRYDOYKQAWDTFCVVVQTK 143
 QY 61 MGEK---VVEFHSIKGGVHRVWIDHPFLAKVWGTGSKLYGPRSGADYLDNHRKPAFLFC 117
 Db 144 VGDKVENVAFHCYKRGVDRFVDHPFIFLAKVVGKTKGSKLYGPRSGADYLDNHRKPAFLFC 203
 QY 118 KAAIEAARVLPF-----GP-GECCVFVANDHWSALVPLVLLKDEYQPKQFTKAKSVLA 169
 Db 204 QAAIEAPQVNLNLSKYSFSGPYGDEVFVANDHWTALLPCYLKSMYQSRGVYVNAKVVFC 263
 QY 170 IINIAPOGRWEEAFKDTKLP---QAADFKLAFSDGYAKVYFTEATPEEDEKPPLTGKTY 226
 Db 264 IINIAPOGRWEEAFKDTKLP---QAADFKLAFSDGYAKVYFTEATPEEDEKPPLTGKTY 303
 QY 227 KINLWLGKGIILADKLVTVPNVAITAAAGGVDELTVIRAKGIEGVINGMDIEEWN 286
 Db 304 RKLNMWKAALIEAHRVLTVPYQAELISGVDRGVELHKYLRMKTYSGLINGMDVQEWNP 363
 QY 287 KTDKFLSAPYDQNSYAGKAAKEALQAEGLGVPDPTAFPLFAFAGRLSEQKGVDIILAL 346
 Db 364 STDKYIDIKYDITVTDAKPLIKEALQAAVGLPVDVDPVVICFICGLEEQKGSILVEAI 423
 QY 347 PKLATPKVQJAILGTGKAAEKLVAIGTKYKGRAGVGVKFSAPLAHMLTAGADPMLVP 406
 Db 424 SKFMGL-NVQMVLGTGKAAEKLVAIGTKYKGRAGVGVKFSAPLAHMLTAGADPMLVP 482
 QY 407 SRFEPGGLIQLHAMHYGTVPVASTGGLVDTVKEGVTGFHMGALNP--DKLDEADADALA 464
 Db 483 SRFEPGGLIQLHAMHYGTVPVASTGGLVDTVKEGVTGFHMGALNP--DKLDEADADALA 542
 QY 465 ATVRASEVFAGGRYPYEMVANCISQDLSWSKPAQKWEGLLEEVYKGVATAKKEIKV 524
 Db 543 KAVTRAVAVYGTSAQEMVKNKMDQDFSNKGPAPLWELKVLNVA-GSEAGTEGEEI-A 600
 QY 525 PVAEK 529
 Db 601 PLAKE 605

RESULT 7
 AAW56484 standard; Protein; 533 AA.
 ID AAW56484;
 AC AAW56484;
 DT 11-SEP-1998 (first entry)
 XX Zea mays waxy gene glucosyl transferase.
 DE SER; starch-encapsulating region; fusion vector.
 KW Zea mays.
 OS Zea mays.
 XX WO9814601-A1.
 PN 09-APR-1998.
 PD 30-SEP-1997; 97WO-US17555.
 PF 30-SEP-1996; 96US-0026855.
 PR (EXSE-) EXSEED GENETICS LLC.
 PA Guan H, Keeling P;
 PI WPI; 1998-240100/21.
 DR N-PSDB; AAV29752.
 XX Hybrid polypeptide comprising starch-encapsulating region and
 PT protein - useful for, e.g. producing protein(s) resistant to
 PT degradation by stomach acids
 XX

Example 2; Page 28; 156pp; English.

The sequence is that of maize glucosyl transferase which is encoded by the waxy gene. It can be used in the production of a hybrid polypeptide comprising a starch-encapsulating region (SER) fused to a payload protein. The hybrid polypeptide can be used to make modified starches comprising the payload protein, selected from, e.g. hormones, growth factors, antibodies, enzymes, dyes, immunoglobulins, etc. The modified starch can also be used to provide grain feeds enriched in amino acids. By encapsulating the payload protein in starch, it is more resistant to degradation by stomach acids.

Sequence 533 AA;

Query Match 40.2%; Score 1353.5; DB 19; Length 533;
Best Local Similarity 50.4%; Pred. No. 5.1e-98;
Matches 275; Conservative 79; Mismatches 153; Indels 39; Gaps 11;

QY	2	LDIVWAAEAVPWSKTTGGLDVGTVGGGLPIELVKGHRVMTIAPRYDQADADAMDTSVWVDI- 60
Db	5	MNVVGAEMAPWSKTTGGLDVGTVGGGLPPAMAANGHRVWVSPRYDQKDAWDTSVSGEIK 64
QY	61	MG---EKVRYFHSIKKGVRHVVHDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRALFC 117
Db	65	MGDGYETVRFFHCYKRGVDVRFVDHPFLFLERVWGKTEEKYGVPGAGTVYDRDQNLRFSLC 124
QY	118	KAATEAARVL-----PF--GP-GEBCVFVANDHSHALVPVLLKDEYQPKGQFTKAKSVLA 169
Db	125	QAALAPRIILSNNNPYFSGPYGEDVVFVCDNHTGFLSCYLKSNYOSHGIYRDAKTAPC 184
QY	170	IHNIAFOGRMWEBAFKDTIKLPQAADFCLAFSDGYAKYITEATPMEDEKPLTGTKYKKI 229
Db	185	IHNISYQGRFAFSDYPENLPNERKSFDEIDGYEK-----PVEG---RKI 227
QY	230	NWLKGGITIAADKLVTYSPNYATETAADAAGVGLDVTIRAKGIEGIVNGMDIEFWNPKT 289
Db	228	NWMAKGLEADRVLTVPYYAEELISGARGCELDNIMRLTGTITGIVNGMDVSEWDSRD 287
QY	290	KFLSAPYDQNSVYAGAAKAEALQAEGLPVDPTAPLFAFTRGLEOKGVDVILIAALPKI 349
Db	288	KYIAVXVDVSTAEAKALKEALQAEVGLPVDNRNIPLVAFTRGLEOKGPDVMAAALPOL 347
QY	350	L-ATPKVQJIAILGTGKAAVEKLVNAIGTKYKRAKGWVKSAPLAHMLTAGADFMLYPSR 408
Db	348	MEMVEDQVILLGTGKKKPERMLMSAEKFPKGVRAVVKENAAALHHIMAGADVLAITSR 407
QY	409	FEPGGLTQLHAMHYGVPPVASTGGLVDYTKGEGTGFHMGALNPD---KLDEADADAJAAT 466
Db	408	FEPGGLTQLOGMYRGTPCACASTGGLVDITIEGKTGFHMGRLSDVCNVVEPADVKKVATT 467
QY	467	VRRASEVAGGRYPPEMVANCISODLSWSKPAQKWGLELLE- -VYVGAGGY-----ATAKK 519
Db	468	LQRAIKVGTTPAEEMVRNMIQDLSWKGPAKWENVLISLGVAGGEPGVEGEIAPLAK 527
QY	520	EEKVP 525
Db	528	ENVAAP 533

RESULT 8
AAY16604

AM16604
ID AAY16604 standard; Protein; 637 AA.

AC AAY16604;

DT 23-AUG-1999 (first entry)

DE Protein encoded by the DNA sequence of the maize waxy gene.

Non-glycogen-like polysaccharide production; fermentation; waxy gene;
 starch synthesis enzyme; glycogen-synthesis enzyme; glycogen synthesis;
 non-starch branching gene; amylopectin; amylose; plant-like starch.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:47:01 ; Search time 16.8268 Seconds
(without alignments)
3719.278 Million cell updates/sec

Title: US-09-980-771a-3_COPY_58_708

Perfect score: 3370
Sequence: 1 ALDIYVVAEVAEPWWSKTGGL.....SASKTSAAPVISAATRKSA 651

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR73: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1481.5	44.0	608	S43341	starch synthase (E
2	1457.5	43.2	607	YUPOY	starch synthase (E
3	1415.5	42.0	603	S61504	glycogen(starch) s
4	1412.5	41.9	608	T10906	starch synthase (E
5	1398.5	41.5	609	JQ0703	glycogen(starch) s
6	1397.5	41.5	609	S11481	glycogen(starch) s
7	1395.5	41.4	603	YUBHY	glycogen(starch) s
8	1390	41.2	610	F86453	granule-bound star
9	1362	40.4	615	1 YWYTY	glycogen(starch) s
10	1355.5	40.2	608	T14731	glycogen(starch) s
11	1353.5	40.2	605	S07314	glycogen(starch) s
12	1194	35.4	238	T07921	glycogen(starch) s
13	902.5	26.8	641	T07668	probable starch sy
14	873.5	25.9	626	JQ2322	starch synthase (E
15	872	25.9	732	T01208	starch synthase (E
16	871	25.8	752	S61505	glycogen(starch) s
17	869	25.8	610	T06280	probable starch sy
18	861.5	25.6	622	T01414	starch synthase (E
19	847.5	25.1	788	T07667	glycogen(starch) s
20	839	24.9	491	T06798	probable starch sy
21	837	24.8	698	T01209	starch synthase (E
22	684.5	20.3	477	B95130	glycogen synthase
23	683.5	20.3	477	H98000	starch synthase (E
24	677.5	20.1	484	S40051	starch synthase (E
25	669	19.9	483	C86712	starch synthase (E
26	652	19.3	480	H98228	glycogen synthase
27	652	19.3	480	AF3057	glycogen synthase
28	651	19.3	486	H72321	glycogen synthase
29	637	18.9	477	A97176	glycogen synthase,

ALIGNMENTS

RESULT 1

S43341

starch synthase (EC 2.4.1.21) precursor - cassava

N;Alternate names: bacterial-glycogen synthase; glycogen synthase; starch synthase
C;Species: Manihot esculenta (cassava)

C;Date: 07-Sep-1994 #sequence_revision 03-Nov-1995 #text_change 16-Aug-2002
C;Accession: S43341

R;Salehuzzaman, S.N.I.M.; Jacobsen, E.; Visser, R.G.F.
Plant Mol. Biol. 23, 947-962, 1993

A;Title: Isolation and characterization of a cDNA encoding granule-bound starch synt

A;Reference number: S43341; MUID:94083565; PMID:8260633

A;Accession: S43341
A;Molecule type: mRNA

A;Residues: 1-608 <SAL>

A;Cross-references: EMBL:X74160; NID:g437041; PIDN:CAA52273.1; PID:g437042

C;Genetics:

A;Gene: GBSS; waxy

A;Genome: nuclear

C;Function:

A;Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producir

C;Superfamily: starch synthase

C;Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltra
F;1-78/Domain: transit peptide (amyloplast) #status predicted <INP>

F;79-608/Product: ADPglucose-starch glucosyltransferase #status predicted <MAT>

Query Match 44.0%; Score 1481.5; DB 2; Length 608;
Best Local Similarity 55.0%; Pred. No. 4.4e-81;
Matches 299; Conservative 64; Mismatches 148; Indels 33; Gaps 9;

QY	2	LDIVMVAEVAEPWWSKTGGLGVDVTGGLP	1	ELVLRKGRHVTIAPRYDQYADAWDTSVVVDI	60
Db	82	MNLIFVGAEGPWSKTGGLGVDVGLG	1	PPAMAARGHRVMTVSPRYDQYKDAWDTSVSVEIK	141
QY	61	MG---EKVYFHSIKKGVHRVMDHFWFLAKVWGKTGSKLYGPRSGADYLDNHKRALPC	117		
Db	142	IGDRIETVRFHFSYKRGVDVDFVDFHMFLEKFWGKTGSKYIYPRAGLDYQDNQLRSLIC	201		
QY	118	KAATGAARVLP-----GP--GECVFFVNDWHSALVPVLKDEYQPKGQFTKAKSVLA	169		
Db	202	LAALGAAPVLNINSSKNFSGPYGEEVAFIANDWHTALLPCYLKATYOPGIYKHAKVAPC	261		
QY	170	INIAFQGMEEAFKDTKLPOAFDKLAFSDGYAKVVTETATPMEEDEKPPLTGTGTYKI	229		
Db	262	INIAVQGRFASDPRLNLPDKFKSSDFIDGYEK-----PVKG---RKI	304		
QY	230	NWLKGITTAADKLIVTSPNYATEIAADAGGVVELDTVIRAKGIEGIVNGMDTEWNP	289		
Db	305	NMKAGILSDRLVITVSPYIAQEVISGVERGVELDNFIRKTKTAGIINGMDYQWNP	364		
QY	290	KFLSAPYDONSYYACGAKEALQAFGLPVDPTAPLPAFIFGLBEQGVDIILALFKI	349		
Db	365	KYIDHYDATVMDAKPILLKEALQAEVGLPDRNVPLTGLFGLRLEQKGSDFVAAISO	423		


```

197 QALAEAPRVNLNLSKSYFSGPYGEDVIFVANDWHSALIPCYILKSMKSYGLYKNAKVAFC 256
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

RESULT 4

Tl0906
starch synthase (EC 2.4.1.21) - sweet potato
N:Alternate names: starch synthase
C:Species: Ipomoea batatas (sweet potato)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Aug-2002
C:Accession: Tl0906
F:Wang, S.J.; Yeh, K.W.; Tsai, C.Y.
submitted to the EMBL Data Library, December 1995
A:Reference number: Z17212
A:Accession: Tl0906
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-608 <WAN>
A:Cross-references: EMBL:U44126; NID:gl172158; PID:gl172159
A:Experimental source: cv. Tainong; tuberous root
C:Genetics:
A:Gene: SS67
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPGlucose producing
A:Pathway: starch biosynthesis
C:Superfamily: starch synthase

Db	262	IHNIAVQGRFAFSFSLNLNLPDEYKGSFDFIDGYDK-----PVKG---RKI	304
QY	230	NWLKGGITTAADKLVTWSPNYATEIADADAGGVELDTVIRAKGIEGIVNGMDTEENWPKTD	289
Db	305	NWKMGIRADRVFTVSPNYAKELVSCVSGVELDNHRCGITGICGMDTQEWNPATD	364
QY	290	KFLSPAPDONSVYAGKAAAEKALQAEGLPLVDPTAPLFAFVIGRLEEQRGVDIILAAALPKI	349
Db	365	KYLAVKYDITITVMOAKPPLLKEALQAAVGLPVDNRNPLIGFTGRLEEQRGSIIYAAISKF	424
QY	350	LATPKVQIAIILGTGKAAAYEKLVNAIGTKYKGRACVVKFSAPLAHMLTAGADFMVLVPSRF	409
Db	425	ISM-DVQIILITGKKKFEQQLQEVNMPDKARGVAKFNPLAHMITAGADFMVLPSRF	483
QY	410	EPCGLIQLHAMHYGTVPVVASTGGGLVDTVKKEGTVGFHMGALNPD--KLDEADADALAATV	467
Db	484	EPCGLIQLHAMRYGTPPCICASTGGGLVDTVKKEGTVGFHMGAFNVDCETVPDVLKVITTV	543
QY	468	RRASEVAGGRYPPEVWANCISQDLSWSKPAOKWGLL-----EENVYKGGVATAJK	519
Db	544	GRALAIYGTLAFTEMIKMSQELSWGPAKNWETVLLSLGVAGSEPGVEGE-ETAPLAK	602
QY	520	EEIKVP 525	
Db	603	ENVATP 608	
RESULT 5			
JQ0703			
glycogen(starch) synthase (EC 2.4.1.11) - rice			
N:Alternate names: granule-bound starch synthase			
C:Species: Oryza sativa (rice)			
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Aug-2002			
C:Accession: JQ0703			
R:Wang, Z.X.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.			
Nucleic Acids Res. 18, 5898, 1990			
A:Title: Nucleotide sequence of rice waxy gene.			
A:Reference number: JQ0703; MUID:91016948; PMID:2216792			
A:Accession: JQ0703			
A:Status: translation not shown			
A:Molecule type: DNA			
A:Residues: 1-609 <WAN>			
A:Cross-references: EMBL:X53694			
A:Experimental source: strain subsp. japonica Hengfeng			
C:Genetics:			
A:Gene: waxy			
A:Introns: 114/1; 141/1; 174/1; 204/1; 225/2; 259/1; 295/3; 377/1; 436/1; 500/1; 529/1			
C:Function:			
A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDP-glucose production			
C:Superfamily: starch synthase			

F:1-72/Domain: transit peptide (amyloplast) #status predicted <TNP>
 F:73-603/Product: UDPglucose-starch glucosyltransferase #status predicted <MAT>

Query Match 41.4%; Score 1395.5; DB 1; Length 603;
 Best Local Similarity 53.3%; Pred. No. 6.1e-76;
 Matches 277; Conservative 71; Mismatches 141; Indels 31; Gaps 7;

```

QY 2 LDIVMAAEVAPWSKGGGLDVTGGPLPIELVKRGHRVMTIAPRYDOYADAWDTSVVVDIM 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 76 MNLVFGAEMAPWSKGGGLDVLGGPLPMAANGHRVMTVPYDQYKDAWDTSVTSEIK 135
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 ----GEKRYFHSIKKGVHRVWIDHPWFLAKVWGKSGKLYGPRSGADYLDNHRKRALFC 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 136 VADEYERVRFFHCYKRGVDVRFIDHPWLEKVRGKKEKIYGPDAAGTYEDNQORSLLC 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 KAAIEAARVL----PF--GDCVFFVANDHWSALVPVLLKDEYOPKQFTKAKSVLA 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 196 QAALEAPRIILNANNPFYSGPYGEDVVFVNCNDWHTGLLACYLKSNGIYRTAKVAF 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 170 IHNIAFQGRWEEAFKDTKLPOAAFDKLAFCSDGYAKVYTEATPMEDEKPPLTGKTKY 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 256 IHNISYGRFSFDFAQLNLPDRFKSFDIDYDK-----PVEG---RKI 298
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 230 NMLKGIIAADKLVTVSPNYATEIAADAAGGVELDTVIRAKGIEGIVGMDIEBWNPKTD 289
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 NMKAGILQADKLVTVSPNYATEIAADAAGGVELDTVIRAKGIEGIVGMDIEBWNPKTD 358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 KFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPFAFVIGRLKDEYOPKQGVDDIILALPKI 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 359 KFLAVNYDITTALEAKALKEALQAEGLPVDPTAPFAFVIGRLKDEYOPKQGVDDIILALPKI 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 350 LATPKVQIALTGTGKAAEKLNAIGTKYKRAKGVVKSAPLAHMLTAGADFMVPSRF 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 419 LKEEDVQIILGTGKPKKPKLLSMEEKPPGKVRVVRNAPLAHMLTAGADFMVPSRF 478
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 410 EPCGLIQLHAMHYGTVPVVASTGGLVDVTKVKGTVGFHMGALNPD--KLDEADADALAATV 467
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 479 EPCGLIQLGMRGYPGVCASGTGGLVDVTKVKGTVGFHMGALNPD--KLDEADADALAATV 538
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 468 RRASEVAGGRYPPEMVANCIQDLSWSKPAQKWEGLLEEV 507
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 539 KRAVKVVGTAPYQEMVKNQMIQDLSWSKPAQKWEGLLEEV 578
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 8

F86453
 granule-bound starch synthase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F86453
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dwyer, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
 C.A.H.; Li, J.H.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maili, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Status: preliminary
 A:Accession: F86453
 A:Molecule type: DNA
 A:Residues: 1-610 <S>TO>
 A:Cross-references: GB:AE005172; NID:G6910568; PIDN:AAF31273.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: starch synthase

Query Match 41.2%; Score 1390; DB 2; Length 610;
 Best Local Similarity 52.7%; Pred. No. 1.3e-75;
 Matches 287; Conservative 72; Mismatches 146; Indels 40; Gaps 12;

```

QY 2 LDIVMAAEVAPWSKGGGLDVTGGPLPIELVKRGHRVMTIAPRYDOYADAWDTSVVVDI- 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 84 MSVIFIGAEVGPWSKGGGLDVLGGPLPAAAGHRVMTICPRVDYQKDAWDTCVVQIK 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 MGEK---VRYFHSIKKGVHRVWIDHPWFLAKVWGKSGKLYGPRSGADYLDNHRKRALFC 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 VGDKVENRFFHCYKRGVDVRFIDHPWLEKVRGKKEKIYGPDAAGTYEDNQORSLLC 203
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 KAAIEAARVLPF-----GP-GEDCVFVANDHWSALVPVLLKDEYOPKQFTKAKSVLA 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 204 QAALEAPQVNLNLSKYSFGPYGEDVVFVANDWHTALLPCYLKSNYQSRGYMNAKVVEFC 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 170 IHNIAFQGRWEEAFKDTKL---QAAFDKLAFCSDGYAKVYTEATPMEDEKPPLTGKTKY 226
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 264 IHNIAFGRAFDYDLSLLNLPISFKSFD---FMDGYEK-----PVKG--- 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 227 KKNILKGGIIAADKLVTVSPNYATEIAADAAGGVELDTVIRAKGIEGIVGMDIEBWNPK 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 304 RKNIMKAAALEAHRLVTVSPNYAQLISGVDRGVLELHKYLRMKTVSGIINGMDVQEWNP 363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 287 KTKFELSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPFAFVIGRLKDEYOPKQGVDDIILAL 346
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 364 STDYKIDIKDIITVDKPLIKEALQAAVGLPVDVPIGIVGRLKDEYOPKQGVDDIILAL 423
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 347 PKLATPKVQIALTGTGKAAEKLNAIGTKYKRAKGVVKSAPLAHMLTAGADFMVPSRF 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 424 SKEMGL-NVQMVILGTGKKEAEQILLEEFKFGKAVGVAKFVNPVLAHMLTAGADFMVPSRF 482
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 407 SRPEPGLQLHAMHYGTVPVVASTGGLVDVTKVKGTVGFHMGALNPD--KLDEADADALA 464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 483 SRPEPGLQLHAMHYGTVPVVASTGGLVDVTKVKGTVGFHMGALNPD--KLDEADADALA 542
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 465 ATVRRASEVAGGRYPPEMVANCIQDLSWSKPAQKWEGLLEEVYKGVATKAKKEIKV 524
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 543 KAVTRAVAYGTSAMQEMVKNQMIQDLSWSKPAQKWEGLLEEVYKGVATKAKKEIKV 600
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 525 PVAEK 529
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 601 PLAKE 605
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 9

YUWYI
 glycogen(starch) synthase (EC 2.4.1.11) precursor - wheat
 N:Alternate names: starch synthase
 C:Species: Triticum aestivum (common wheat)
 C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002
 C:Accession: S16261; S33636
 R:Clark, J.R.; Robertson, M.; Ainsworth, C.C.
 Plant Mol. Biol. 16, 1099-1101, 1991
 A:Title: Nucleotide sequence of a wheat (Triticum aestivum L.) cDNA clone encoding the
 A:Reference number: S16261; MUID:9122506; PMID:1863765
 A:Accession: S16261
 A:Molecule type: mRNA
 A:Residues: 1-615 <CLA>
 A:Cross-references: EMBL:X57233; NID:g21901; PIDN:CAA40509.1; PID:g21902
 R:Ainsworth, C.; Clark, J.; Baisdon, J.
 Plant Mol. Biol. 22, 67-82, 1993
 A:Title: Expression, organisation and structure of the genes encoding the waxy protei
 A:Reference number: S33636; MUID:93271462; PMID:8499619
 A:Accession: S33636
 A:Molecule type: protein
 A:Residues: 71-75, 'X', 77-78 <AIN>
 C:Genetics:
 A:Gene: waxy
 C:Function:
 A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose produc
 A:Pathway: starch biosynthesis
 C:Superfamily: starch synthase
 C:Keywords: amyloplast; glycogen/starch biosynthesis; glucosyltransferase; hexosyltra
 F:1-70/Domain: transit peptide (amyloplast) #status predicted <TNP>
 F:71-615/Product: UDPglucose-starch glucosyltransferase #status experimental <MAT>

Query Match	40.2%;	Score 1355.5;	DB 2;	Length 608;
Best Local Similarity	50.6%;	Pred. No. 1.5e-73;		
Matches 276: Conservative	76;	Mismatches 154;	Indels 39;	Gaps 11;

RESOUT 11
S07314

glycogen (starch) + ornithine

C; Date: 10

R; Kloesgen
Mol Gen

A; Referen

A; Molecule

A; Experiment

A; Note: p

A: Introns

A;Descrip

F: 1-72/DO

Best To
Matches